Proteomic profile of cystic fibrosis sputum cells in adults chronically infected with Pseudomonas aeruginosa

EUROPEAN RESPIRATORY journal
OFFICIAL SCIENTIFIC JOURNAL OF THE ERS

Proteomic profile of cystic fibrosis sputum cells in adults chronically infected with *Pseudomonas aeruginosa*

Journal:	*European Respiratory Journal*
Manuscript ID	ERJ-01569-2016.R1
Manuscript Type:	Original Article
Date Submitted by the Author:	15-Mar-2017
Complete List of Authors:	Pattison, Sally; Queen's University Belfast, Centre for Infection and Immunity Gibson, David; Queen's University Belfast, Centre for Infection and Immunity Johnston, E; Queen's University Belfast, Centre for Infection and Immunity Peacock, Samantha; Cold Spring Harbor Laboratory, Pappin Laboratory Rivera, Keith; Cold Spring Harbor Laboratory, Pappin Laboratory tunney, michael; Queen's University Belfast, School of Pharmacy Pappin, Darryl; Cold Spring Harbor Laboratory, Pappin Laboratory Elborn, J; Queen's University Belfast, Centre for Infection and Immunity
Key Words:	cystic fibrosis and chronic infections, sputum, proteomics

Reviewer: 1

Comments to the Author –
This is an interesting paper using a novel -omics Technology to study CF Airway inflammation.
While the approach, the originality and Interpretation of the findings are sound and of high quality, there Are several issues that should Be Adressed:

Reviewer 1 Comment 1: the number of included CF Patients is low, which is understandable given the omics approach. However, As ERJ is a clinically-focused rather than a biochemical Journal, this part should be strenghtened, for instance: (i) include additional CF Patient samples, (ii) add some longitudinal analyses, (iii) add a disease control Group for instance COPD or BE, (iv) add some BAL samples (If available) to validate their findings in sputum.

Reviewer 1 Response 1: *Our study was a clinically-focused investigation of the potential mechanisms for lung damage in adults with cystic fibrosis. It uniquely studied the protein abundance of cells in vivo, thereby taking into account both their natural frequency in the sputum cell population and the effects of intercellular interactions. Thus, it delivers a high degree of clinical relevance to our findings. The manuscript has been revised to include a more detailed analysis of longitudinal samples for each of the CF patients enrolled in the study. Furthermore, we have analysed the changes in sputum total protein yield (mg per g sputum) and differential relative abundance in individual proteins of interest during intravenous antibiotic treatment of pulmonary exacerbation. In both cases we observed that antibiotic treatment of exacerbation is associated with proteomic changes to more closely resemble the healthy control cohort. We have also observed that at the follow-up sample collection total protein yield returned to levels similar to exacerbation pre-treatment levels.*
Given cost and time restrictions, we have focused exclusively on CF. However, if further funding were to come available, we would like to expand our investigation to include other cohorts.

Reviewer 1 Comment 2: The authors identified interesting correlations for AAT and Olfactomedin-4 in their small CF cohort. The authors could go on and validate the broader clinical usefulness of These markers in a larger CF patient cohort.

Reviewer 1 Response 2: *While we agree that it would be extremely interesting to follow-up on candidate biomarkers identified by our study, such as alpha-1-antitrypsin and olfactomedin-4, the scope of our study was exploratory. We aimed to investigate for the first time the global proteome of the combined population of cells present in CF sputum in order to assess their potential hazard through cell activities and/or release into sputum fluid. We have highlighted the exploratory nature of the study and our acknowledgement of this limitation has generally been recognised by the reviewers in light of the novel proteomics approach both in terms of technology and the previously unstudied cell population that more accurately represents the CF lung than previous studies of single cell type in vitro cultures or animal models. We do not currently have ethical approval and funding to expand its scope or for the increased patient recruitment which would be required in order to investigate a larger CF patient cohort. However, in response to reviewer's comments we have undertaken additional laboratory analysis using our existing sample set to validate findings using commercial ELISAs and so have demonstrated that commercial ELISAs are a practical high-throughput method suitable for independent, large-scale validation of candidate biomarkers identified by this study.*

Reviewer 1 Minor points –

Reviewer 1 Comment 3: introduction and Discussion could include more Refs on neutrophils in CF

Reviewer 1 Response 3: *While we agree that neutrophils play an important role in CF, our study investigated all cell types found in CF sputum and we feel that to comment specifically on neutrophils would require us to comment on all cells likely to be present in our samples. Unfortunately we are unable to do this due to journal requirements that restrict the number of references. Also, we wished to avoid lengthening the manuscript.*

Reviewer 1 Comment 4: Ref 38 should Be substituted, for instance with Manzenreiter et al JCF and/or Marcos et al. Mediators of Inflammation 2015.

Reviewer1 Response 4: *Thank you for your suggestion. We have now substituted the following reference-*

Marcos V, Zhou Z, Yildirim AO, Bohla A, Hector A, Vitkov L, Krautgartner WD, Stoiber W, Griese M, Eickelberg O, Mall MA, Hartl D. Free DNA in cystic fibrosis airway fluids correlates with airflow obstruction. Mediators Inflamm 2015; 2015:408935.

Reviewer: 2

Comments to the Author

This is a proteomic study of 12 adult patients with cystic fibrosis. I enjoyed reading the paper, and it is likely to be very interesting as a hypothesis generating manuscript for those searching for CF biomarkers. The paper is very nicely written and easy to follow.

Comments

Reviewer 2 Comment 1: Early in the manuscript (perhaps even in the title) and in some parts of the discussion, it would be nice to be clear that the conclusions in the paper are relevant to "adult CF patients with P. aeruginosa colonisation". The proteome results here cannot be used to make conclusions about "CF" but rather about "adult CF patients with P. aeruginosa infection".

Reviewer 2 Response 1: *We have amended the title of the manuscript to "Proteomic profile of cystic fibrosis sputum cells in adults chronically infected with Pseudomonas aeruginosa". In addition we have now recognised this study limitation more consistently throughout the manuscript.*

Reviewer 2 Comment 2: Did the authors perform any comparisons of proteomic data between stability and exacerbation?

Reviewer 2 Response 2: *The manuscript has been revised to include a more detailed analysis of longitudinal samples collected over the course of intravenous antibiotic treatment of pulmonary exacerbation. Accordingly, a new results section entitled "Longitudinal analysis of antibiotic treatment of exacerbation" has been added and the findings commented on in the discussion.*

Reviewer 2 Comment 3: How do the authors differentiate between NET proteins and those that are simply released from neutrophils during cell death or lysis during the protein extraction process? I accept that prior studies have found NETS in the CF lung, but it seems speculative to say these proteins are NETs (unless the authors have some data to show that specific neutrophil proteins are increased relative to other neutrophil proteins that are not implicated in NETs). This issue also leads into point 4

Reviewer 2 Response 3: *From our data we cannot definitively say that any of the proteins detected were NET proteins and indeed we have explicitly stated in our manuscript discussion (P16) that as "our study focused on sputum cellular proteins we did not look specifically for NET structures, which are extracellular, and therefore do not know if they were present." Instead, what we have reported is the consistently high abundance in sputum cells of proteins which are known to be the main constituents of NETs. Although we cannot know the fate of these cellular proteins, NETs are a recognised phenomenon in CF and we think it is reasonable to suggest that they could pose a significant hazard if indeed they were released as NETs. However, in order to avoid misinterpretation of our discussion of this potential hazard we have revised wording where appropriate throughout the manuscript.*

Reviewer 2 Comment 4: It is a shame that some of the candidates were not validated by ELISA, but the authors have acknowledged this issue. None of the identified candidates can be considered biomarkers until validated. I think largely the authors have largely addressed this well (except for the speculation about NETs above).

Reviewer 2 Response 4: *As explained in Response 2 to Reviewer 1, we agree that it would be extremely interesting to follow-up on candidate biomarkers identified by our study in a large independent cohort. As a first step towards such validation, we have performed and detailed in the manuscript additional laboratory analysis of our existing sample set to confirm findings for two candidate biomarkers, myeloperoxidase and lactotransferrin. Thus, we have demonstrated the potential of commercial ELISAs for further validation of candidate biomarkers in an independent large-scale study. We have highlighted the exploratory nature of our study and our acknowledgement of this limitation has been generally recognised by the reviewers in light of the novel proteomics approach both in terms of technology and the previously unstudied cell population.*

Reviewer 2 Comment 5: I believe reference 38 has been retracted. It is bad practice to reference papers that have been retracted.

Reviewer 2 Response 5: *We have now substituted the following reference –*

Marcos V, Zhou Z, Yildirim AO, Bohla A, Hector A, Vitkov L, Krautgartner WD, Stoiber W, Griese M, Eickelberg O, Mall MA, Hartl D. Free DNA in cystic fibrosis airway fluids correlates with airflow obstruction. Mediators Inflamm 2015; 2015:408935.

Reviewer 2 Comment 6: The discussion of the negative effects of alpha-1 and olfactomedin 4 in CF in the discussion are a bit simplistic - the reality of inflammation is that both pro and anti-inflammatory molecules are released concurrently during inflammation, and so the implication the authors make that because they are associated with lung function they are likely to be harmful might be true, but it is equally possible these are "innocent bystanders" or an ineffective counter-response. (when the house is on fire, the equivalent of proteomics might detect fire-trucks, but it doesn't mean the fire trucks are causing the fire.....)

Reviewer 2 Response 6: *We agree that a negative correlation with lung function does not prove causation and have revised the manuscript to clarify this.*

Reviewer 2 Comment 7: The introduction is relatively long and i feel some of the last couple of paragraphs could be moved to the discussion. As a minor point, ERJ does not require the acknowledgement of abstract publications (this is a AJRCCM requirement) so this could be removed.

Reviewer 2 Response 7: *We have shortened the manuscript's introduction as suggested.*

Reviewer: 3

Comments to the Author

The information about the global proteome of cells present in CF sputum may integrate the data concerning the extracellular proteome of this fluid. The present manuscript appears as a preliminary report describing the proteome analysis of sputum cells from these patients. As such, the identification of the most abundant cellular proteins is indeed of interest for gaining further understanding of the mechanisms underlying this severe lung disease.
In my opinion this is the novelty of the present manuscript that, in other respects, follows the "conventional" procedure of several similar articles: identification of a large list of proteins, some of which are indicated as candidate biomarkers for disease status. A great deal of future research will obviously be needed to validate these data. Thus, while not having impressed me that much, this article is well-written and science is good.
Nevertheless, I have two major concerns:

Reviewer 3 Comment 1: The major limitation of this study is represented by the very small cohort of individuals analyzed, also in the light of the dishomogeneity of patients considered. As shown in the demographic table, two of them have very high FEV1 values and two very low. If the FEV1 values reflect the severity of the disease (i.e. lung conditions), then one can imagine that the "starting" conditions of these people are not the same. Most likely also in terms of cellular proteome. The authors state that "Lung function was measured as FEV1 % predicted prior to each sputum collection. This was done to encompass all aspects of CF lung health and prevent proteome bias towards any particular clinical status" but what does this mean ? This sentence is ambiguous. It should be re-formulated and moved at the beginning of the Results section indicating the reasons why these patients have formed a "group". However, I wonder whether these data may be actually representative of the proteome of cells present in CF sputum.

Reviewer 3 Response 1: *The inclusion criteria for our study limited the CF cohort to adults with Pseudomonas aeruginosa infection. We agree that other criteria could also have been applied to specify the cohort further, but this would have further limited the clinical relevance of our findings. We aimed to investigate more broadly the sputum cellular proteome of adults with CF chronically infected with P. aeruginosa. Therefore, not only do the FEV_1 values of the CF cohort vary, but we have also deliberately included in our analysis samples collected at different times of clinical health for each individual. The inclusion of this range of samples is what we are referring to with the statement "This was done to encompass all aspects of CF lung health and prevent proteome bias towards any particular clinical status." As requested we have clarified this further at the start of the results section of the revised manuscript.*

Reviewer 3 Comment 2: Given that the number of males and females in the patient's cohort is almost the same, it is not clear to me the reason why the control group is made by a number of females that is two-fold higher than that of males. I think that finding a number of healthy males/females comparable to that of patients was likely not such a big problem. While being aware that the real discrimination is between the proteome of healthy and CF patients, I wonder whether this gender difference may not introduce an additional variable.

Reviewer 3 Response 2: *We agree complete age and sex matching between control and patient groups would be desirable but there is no evidence here of a sex bias in the data for either group.*

Reviewer: 4

Comments to the Author

In this paper, Pattison et al presents a Multidimensional Protein Identification Technology that is applied to analyse sputum cellular proteins from Cystic fibrosis patients in comparison to healthy controls. By utilizing Ingenuity Pathway Analysis, Gene Ontology functions, protein abundance and correlation, claims are made on clinical significance.of the candidate proteins identified in the study.

There are a number of unclear parts within this study that makes the results and claims made questionable. The n-numbers of the patients are really low in order to reach statistical significance here, and it's a major issue with the study. Additional considerations needs to be made before the paper can be considered for further Editorial processing, and they are as follows;

Reviewer 4 Comment 1: How was the samples normalized. Usually the sputum samples are highly heterogeneous, what the status in this study. Already when you mention plugs it's a challenge to get similar samples to be comparable

Reviewer 4 Response 1 (N.B. Since the second point above expands on the concern about sample heterogeneity raised in the first point we have combined our response): *As with all sputum samples, the samples collected for this study were heterogeneous. However, our study followed standardised procedures for the processing of all samples including processing by the same researcher under the same conditions using the same equipment in order to minimise technical variation. Additionally, an equal amount of protein per sample (70 µg) was analysed by mass spectrometry and our data analyses examined the relative protein composition of sputum cells which a lower cell density in one plug compared to another would not affect.*

Reviewer 4 Comment 2: How did you manage to solubilize the samples. After experimental prepp. steps, what was the non-soluble fraction of each and every sputum sample ?

Reviewer 4 Response 2: *As outlined in the Methods section of the Supplementary Data the cells were lysed in a buffer consisting of 10 mM HEPES, 0.5 mM EDTA, 1% v/v Nonidet P 40, 4% w/v SDS, pH 8.0, which also served to solubilise the proteins. We did not analyse the non-soluble fraction.*

Reviewer 4 Comment 3: How did you build a reliable sample preparation methodology – here? I miss out on Recovery data, reproducibility and rigidity of the sample preparation is crucial. If this step is not solid, the entire Expression study is questionable.

Reviewer 4 Response 3: *As already indicated in Response 1 to Reviewer 4, our protocols were standardised and followed rigidly to reduce technical variability. As requested, the total protein yield for each sample (mg per g of sputum) has now been added to the revised manuscript in Figure E1 in the online supplement. We agree that consistent sample preparation is key to expression studies and believe that Fig. 1, which shows that protein profiles tended to cluster by CF individual despite sample collection and processing on separate days, demonstrates that reproducibility was achieved in our study. Further evidence of method reproducibility is provided by the consistent pattern for each CF participant of a decrease in protein yield (mg per g of sputum) after antibiotic treatment of exacerbation, followed by a return to higher levels in the stable sample for 11/12 participants, as shown in Figure E1.*

Reviewer 4 Comment 4: How did you normalize the samples How did you manage to correlate the patient groups here to make them comparable in the study.

Reviewer 4 Response 4 (since the above points are linked we have made one response): *As described in the Methods section of the Supplementary Data, 70ug of soluble protein was trypsin-digested and iTRAQ labelled per sample and six samples plus a reference were combined in each mass spectrometry run. Each run's internal reference was then used to normalise the samples to obtain a relative abundance for each protein (relative to the reference) per sample. Normalisation to the reference sample allowed us to compare between cohorts. This process has been clarified in the Statistical analyses section.*

Reviewer 4 Comment 5: Another remark by this reviewer is that the numbers of patient samples are really low in order to make any conclusions….

Reviewer 4 Response 5: *We have explicitly acknowledged in our manuscript that our study has a relatively low sample size and explained that this is the inevitable trade-off for a quantitative, high proteome coverage analysis of sputum cells as a novel target. We have additionally explained in the manuscript that while cohort size limits the scope of our study conclusions, the findings are of exploratory interest and may have future clinical value if validated in a larger study. In general, the reviewers have accepted our justification of this limitation to our study.*

Reviewer 4 Comment 6: I think its unacceptable in 2016 to present page don and page up of annotated and regulated proteins from 12 patients when you have no Validation, except for 2 WB experiments…..

- Further to the above, you state on page 5;

Differentially abundant proteins related to Rho family small GTPase activity, immune cell movement/activation, generation of reactive oxygen species and the dysregulation of cell death and proliferation. Compositional breakdown established neutrophil extracellular trap associated proteins as consistently detected in high abundance. A further 17 proteins correlated negatively with lung function".

Further, you continue;

These findings expand current understanding of the mechanisms underlying inflammatory drivers in CF lung disease and identify sputum cellular proteins with potential for use as markers of disease status, prognostic indicators, stratification determinants for treatment prescription or as therapeutic targets."

This reviewer cannot find data that supports the valiudity of the se claims made here. It seems that to fulfil these, another much bigger validation study needs to be put in place to support the data in this study.

- The claims made are simply not possible to prove with statistical significance
- I would like the authors to have a view on the protein regulation and the disease biology …

Reviewer 4 Response 6 (as the above points are linked we have made one response): *The discovered proteins indeed act as a set of candidates which we agree requires a much larger and correctly powered follow-up validation study. In response to reviewer's comments, we have undertaken additional laboratory work and validated the findings in our existing cohort using commercial ELISAs. Moreover, we have demonstrated that commercial ELISAs are a practical high-throughput method suitable for future independent, large-scale validation of candidate biomarkers identified by our study.*

Proteomic profile of cystic fibrosis sputum cells in adults chronically infected with *Pseudomonas aeruginosa*

Sally H. Pattison,[1] David S. Gibson,[1] Elinor Johnston,[1] Samantha Peacock,[2] Keith Rivera,[2] Michael M. Tunney,[3] Darryl J. Pappin,[2] and J. Stuart Elborn[1]

[1] Centre for Infection and Immunity, Queen's University Belfast, Belfast, Northern Ireland, BT9 7BL, UK.

[2] Cold Spring Harbor Laboratory, Cold Spring Harbor, NY 11724, USA.

[3] School of Pharmacy, Queen's University Belfast, Belfast, Northern Ireland, BT9 7BL, UK.

Corresponding Author: Dr Sally Pattison, Centre for Infection and Immunity, Medical Biology Centre, 97 Lisburn Road, Queen's University Belfast, N. Ireland, BT9 7BL, UK. Email: s.pattison@qub.ac.uk. Phone: (+44) 028 9097 2698

Take-Home Message: CF sputum proteomics detects high abundance of NET proteins and identifies proteins correlating negatively with $\%FEV_1$.

ABSTRACT

Lung disease is the main cause of morbidity and mortality in cystic fibrosis (CF) and involves chronic infection and perturbed immune responses. Tissue damage is mediated mostly by extracellular proteases, but other cellular proteins may also contribute to damage through their effect on cell activities and/or release into sputum fluid by means of active secretion or cell death.

We employed Multidimensional Protein Identification Technology to identify sputum cellular proteins with consistently altered abundance in adults with CF, chronically infected with Pseudomonas aeruginosa, compared to healthy controls. Ingenuity Pathway Analysis, Gene Ontology, protein abundance and correlation with lung function were used to infer their potential clinical significance.

Differentially abundant proteins relate to Rho family small GTPase activity, immune cell movement/activation, generation of reactive oxygen species and dysregulation of cell death and proliferation. Compositional breakdown identified high abundance of proteins previously associated with neutrophil extracellular traps. Furthermore, negative correlations with lung function were detected for 17 proteins, many of which have previously been associated with lung injury.

These findings expand current understanding of the mechanisms driving CF lung disease and identify sputum cellular proteins with potential for use as indicators of disease status/prognosis, stratification determinants for treatment prescription or as therapeutic targets.

INTRODUCTION

The lack or reduced presence at cell surfaces of fully functional cystic fibrosis conductance regulator (CFTR) in people with cystic fibrosis (CF) results in a build-up of dense, dehydrated mucus that promotes infection and thus, chronic inflammation [1]. Additionally, CF has been labeled a mucosal immunodeficiency syndrome because of the direct proinflammatory effects of CFTR dysfunction on the dysregulation of epithelial innate immunity and airway leukocytes [2]. As such, the pathophysiology of the CF lung is highly complex, involving numerous interactions between multiple human cell types and a diverse microbiota [3-5], as well as the inherent impairment of several cell types.

Proteomic studies of the CF airways have investigated protein abundance in epithelial cells cultured *in vitro* or obtained from animal models or clinical samples [6-13]. Furthermore, several studies have assessed the extracellular proteomes of CF sputum, bronchoalveolar lavage fluid (BALF) and bronchial epithelial cell secretions [14-19]. The main observations of these studies are summarised in the online supplement (Table E1). Although these studies have all contributed to our understanding of CF, investigations of single cell types, *in vitro* cultures and animal models have limited relevance to *in situ* CF lung pathogenesis. To date, the contribution of the global proteome of all cell types in the proportions occurring naturally in CF sputum has not been investigated. We hypothesized that proteins detected in sputum cells can have a detrimental effect on CF lung health, both through their effect on cell functioning and activities, and through their release into extracellular sputum fluid by means of active secretion or cell death.

We employed quantitative comparative proteomics to characterise the global activity and protein composition of cells in CF sputum from adults chronically infected with *P.*

aeruginosa to provide insight into potential mechanisms of lung disease and to identify candidate biomarkers associated with poor lung function in this population. By employing these methods we were able to identify, in clinical samples, novel cellular proteins and activities likely to be clinically relevant as mechanisms for tissue damage and disease progression.

METHODS

Study participants

Twelve adults with CF, aged 26.8 ± 5.98 (mean ± SD; range 18-37) were recruited at Belfast City Hospital on admission for treatment with intravenous antibiotics for an acute pulmonary exacerbation as defined by Fuchs et al. [21]. Seven were male and five female. All were chronically infected with *P. aeruginosa*. Their characteristics and treatment regimens are recorded in Table 1 and Table E2 (online supplement), respectively. Three induced sputum samples were collected from each CF participant at times of differing clinical status – untreated pulmonary exacerbation, within 24h of completing intravenous antibiotic treatment of exacerbation and 4-15 weeks after completion of treatment. Lung function was measured as FEV_1 % predicted prior to each sputum collection. Twelve non-smoking, healthy control participants, aged 29.6 ± 4.27 (mean ± SD; range 25-39) were recruited. Eight were female and four male. A single induced sputum sample was collected from each control individual. Ethical approval for this study was granted by the Office for Research Ethics Committees Northern Ireland (Ref: 09/NIR02/68) and informed consent was obtained from all subjects.

	Age	M/F	PEx (previous 12 months)	% FEV_1	CFTR Genotype
CF1	23	F	2	35-36	F508del/ F508del
CF2	36	F	0	26-36	F508del/ F508del
CF3	31	M	1	81-102	F508del/G551D
CF4	18	M	1	57-69	F508del/ F508del
CF5	25	M	2	52-60	F508del/ F508del
CF6	37	F	0	51-62	F508del/ F508del
CF7	28	M	0	26-30	F508del/ F508del
CF11	29	F	1	49-75	F508del/ F508del
CF12	22	M	2	31-41	F508del/ F508del
CF13	19	M	1	55-77	F508del/ F508del
CF15	28	F	1	73-98	F508del/R117H
CF16	26	M	1	40-54	F508del/ F508del

Table 1: Characteristics of the CF cohort. % FEV_1 records the range measured over the 3 sputum collections.

Sample preparation

For each sputum sample the cell population was harvested from mucus plugs, washed and proteins extracted. Briefly, approximately 1g of mucus plug was homogenised thoroughly with 9 ml sterile phosphate buffered saline and the cells pelleted by centrifugation. Following two further wash steps and cell straining to remove clumps, the cells were pelleted to remove all traces of supernatant. The pellet was then homogenised thoroughly in lysis buffer and subjected to repeated freeze-thaw and water bath sonication to ensure cell lysis. Insoluble debris was removed by centrifugation. All work was carried out at 4^0C and in the presence of a protease inhibitor cocktail. Full method details are given in the online supplement.

Proteomic methods

Protein digestion and iTRAQ labelling were performed as described previously {{904 Scuoppo,C. 2012}} and details are recorded in the online data supplement. Protein digestion with trypsin and labelling with 8plex iTRAQ Reagents (AB Sciex) were performed as described previously [22]. Six test samples and a reference sample were mixed for simultaneous analysis by 2D liquid chromatography MudPIT on line with an LTQ Orbitrap Velos (Thermo Electron) as described previously [22]. Further details of the methods employed for MudPIT are recorded in the online data supplement.

Protein identification

Peaklists were generated by Mascot Distiller and protein identification was performed by the Mascot search engine against the UniProt database plus NCBI non-redundant database, taxonomy *P. aeruginosa*. Further details of the methods used in protein identification are recorded in the online data supplement. The mass spectrometry proteomics data have been deposited to the ProteomeXchange

Consortium [23] via the PRIDE partner repository with the dataset identifier PXD001985 and 10.6019/PXD001985. Western immunoblot and enzyme-linked immunosorbent assay (ELISA) were used to confirm the mass spectrometry identification and quantification data as described in the online data supplement.

Statistical analyses

SPSS was used for all statistical analysis unless otherwise stated. iTRAQ sample labels were used to calculate each protein's abundance per sample relative to the reference included in each mass spectrometry run, thus enabling comparison of protein abundance between samples. Hierarchical clustering of sample protein profiles was performed by PermutMatrix [24] using Euclidean distance dissimilarity and McQuitty's linkage criteria. Prior to cluster analysis proteins within the dataset were tested for collinearity (Bivariate Spearman correlation coefficient >0.9) and four proteins that correlated highly with others were subsequently excluded. Independent samples T-tests and Mann Whitney U tests were performed to identify differentially abundant proteins (p<0.05) between the CF and control cohorts. Ingenuity Pathway Analysis (IPA) was then employed to predict: (a) Canonical Pathway activity in both the CF and control consensus core proteomes (Fisher's Exact Test); (b) altered Bio Function activity in CF relative to the control group (Fisher's Exact Test); (c) Bio Function up- or down-regulation in CF (IPA Regulation z-score algorithm).

For proteins detected in >80% of CF samples, their percentage abundance within each sample was estimated by the percentage exponentially modified Protein Abundance Index (% emPAI [25]) using MS Excel. For proteins of consistently high abundance (>1% emPAI) in CF, % emPAIs were compared between the CF and control cohorts (Mann-Whitney U test).

Bivariate Spearman correlation analysis and univariate analysis were performed to identify proteins whose % emPAIs correlated with FEV_1 % predicted and to identify independent factors. A Benjamini-Hochberg multiple comparisons adjustment was performed for the correlation analysis (MS Excel) to control the false discovery rate at <0.05. For each of the proteins correlating negatively with FEV_1, paired T-tests and Wilcoxon signed-ranks tests were performed to compare the relative abundance (sample:reference) of matched CF pre-treatment and <24h post-treatment samples to identify differential abundances associated with antibiotic treatment of exacerbation. Again, the false discovery rate was controlled at <0.05 by a Benjamini-Hochberg multiple comparisons adjustment.

RESULTS

Core signature proteome in CF cohort

In accordance with inclusion criteria, the CF cohort comprised adults with CF who were chronically infected with *P. aeruginosa*. The cohort's variation in % FEV_1 predicted (26-102%) and collection of 3 sputum samples from each individual (covering clinical stability, exacerbation and following antibiotic treatment of exacerbation) prevented proteome bias towards any particular clinical status within this population. In total, 2210 human proteins were detected including 119 which were common to all 48 samples, including those from healthy controls. A further 38 proteins were detected in all samples belonging to only one cohort suggesting low abundance or absence in the other cohort, although differential abundance between the two cohorts cannot be proven statistically. Twenty-one proteins were detected only in the CF cohort and 17 proteins were detected only in the control cohort. Table

E3 in the online data supplement records all proteins detected in all the samples belonging to at least one cohort.

The heatmap in Figure 1, which clusters individual sputum samples according to relative abundance levels (sample:reference) of consistently detected proteins, shows that CF and control samples largely cluster separately. The two exceptions, CF3.2 and CF4.2 (marked with an arrow in Figure 1), are CF samples collected just after completion of antibiotic treatment for a pulmonary exacerbation.

IPA Canonical Pathways in CF and control cohorts

The 136 proteins detected in all control samples and the 140 proteins detected in all CF samples were treated as consensus core proteomes for the two cohorts. Qualitative individual analysis of these proteomes was performed by Ingenuity Pathway Analysis (IPA) to determine Canonical Pathways predicted by IPA as active in each cohort. Figure 2 shows the top 5 Pathways predicted in the control proteome and all Canonical Pathways predicted by to have a >5-fold increase in *p*-value in the CF proteome compared to the control proteome, indicating likely up-regulation in CF. Those predicted as most up-regulated in CF are involved in immune functions including neutrophil recruitment, rearrangement of the actin cytoskeleton, phagocytosis and T cell signaling.

IPA Bio Function analysis of the CF differential proteome

Comparison of relative abundance levels (sample:reference) between the two cohorts (36 CF samples versus 12 control samples) was performed for the 119 proteins detected in all samples in order to identify cohort differences. 54 proteins were found to be less abundant and 36 more abundant in the CF versus control cohort (independent samples T-tests & Mann Whitney U tests, $p<0.05$; Table E3, online data supplement). Abundance ratios (CF:control) for all differentially abundant

proteins were included for IPA Bio Function analysis irrespective of the degree of difference. Table E4 (online supplement) records the Bio Functions predicted to be differentially expressed (Fisher's Exact Test, $p<0.05$) in CF with Regulation z-scores >1 or <-1 (IPA Regulation z-score algorithm, $p<0.05$), and those predicted with a higher confidence level ($p<0.0005$), but with a less noteworthy or absent Regulation z-score (shaded rows).

Inflammatory Response, Cellular Movement, Cell Death, Cellular Growth and Proliferation, and Free Radical Scavenging were predicted by IPA as most altered in CF compared to the control cohort, with regard to the extent of regulation (Regulation z-score). This is largely supported by the data in Table E5 (online data supplement), which records Gene Ontology (GO) biological process classifications for the 38 proteins detected consistently in only one cohort, and the 6 proteins identified by Mann-Whitney U test as showing differential abundance between the two cohorts.

High abundance proteins in CF sputum cells

As high abundance proteins have greater potential for directly impacting the lung environment, we identified proteins that were consistently detected at high abundance per sample in the CF cohort and determined if their abundance levels per sample differed significantly from those of the control cohort. Six proteins were found to each comprise >1% of total human cellular protein per sample (% emPAI) in >80% of CF samples. Comparison of the % emPAIs for these proteins between CF and control cohorts detected increased abundance of all six in CF (independent samples Mann-Whitney U test, $p<0.01$). Their median±IQR % emPAIs are displayed in Figure 4 along with the corresponding data for the control cohort. The Human Protein Atlas [26] identifies all as being either highly or moderately expressed in bone marrow

hematopoietic cells (50% of which are myelopoietic cells) and all except cathepsin G as moderately or lowly expressed in lung macrophages. No statistically significant difference in relative abundance level (sample:reference) for these 6 proteins was detected when comparing paired CF pulmonary exacerbation samples before and <24h after antibiotic treatment (Wilcoxon signed-ranks test, $p<0.05$).

Sputum cellular proteins correlate with lung function

In order to identify any associations between specific proteins and lung function, the relative abundance levels (sample:reference) of the 337 proteins detected in >80% CF samples were tested for correlation with lung function as indicated by FEV_1 % predicted. Negative correlations (Bivariate Spearman correlation analysis, $p<0.01$; false discovery rate <0.05) were identified for the 17 proteins listed in Table 2, implicating them as potential biomarkers of CF lung disease. Positive correlations were also identified, but as the focus of this analysis was to identify potentially injurious proteins only the negative correlations are presented here. Alpha-1-antitrypsin and olfactomedin-4 were independent factors for FEV_1 % predicted (univariate analysis, $p<0.01$).

UniProt Accession	Description	Gene name	Spearman correlation (FEV_1)	Signif. (*p*-value)	CF pre-/post-abx
D3DSM0	Integrin beta-2 *	ITGB2	-.751	0.00000	1.114
P80723	Brain acid soluble protein 1 *	BASP1	-.732	0.00000	1.343
P02675	Fibrinogen beta chain **	FGB	-.732	0.00000	1.690
P01009	Alpha-1-antitrypsin	SERPINA1	-.674	0.00001	
P30048	Thioredoxin-dependent peroxide reductase, mitochondrial	PRDX3	-.646	0.00012	
J3QLC9	Haptoglobin *	HP	-.617	0.00028	1.288
P08246	Neutrophil elastase	ELANE	-.608	0.00036	
Q08722	Leukocyte surface antigen CD47 *	CD47	-.608	0.00037	1.115
Q6UX06	Olfactomedin-4 **	OLFM4	-.601	0.00011	1.620
P15144	Aminopeptidase N	ANPEP	-.540	0.00067	
F8VV56	CD63 antigen *	CD63	-.586	0.00068	1.338
C9JNR4	Transforming protein RhoA	RHOA	-.578	0.00082	1.354
P01023	Alpha-2-macroglobulin	A2M	-.543	0.00192	
O95498	Vanin 2 *	VNN2	-.540	0.00205	1.475
O94804	Serine/threonine-protein kinase 10	STK10	-.535	0.00233	
E7ER45	Maltase-glucoamylase, intestinal **	MGAM	-.523	0.00106	1.314
P08473	Neprilysin	MME	-.511	0.00146	

Table 2: Correlation between relative abundance level (sample:reference) and lung function (FEV_1 % predicted) for proteins present in >80% CF samples at false discovery rate <0.05. * and ** indicate proteins with significantly different abundance

levels (p<0.05 and p<0.01, respectively) in the CF cohort when comparing matched pre-antibiotic and <24h post-antibiotic treatment samples. Median ratios (pre-treatment : <24h post-treatment) are recorded in the final column.

Longitudinal analysis of antibiotic treatment of CF exacerbation

Figure E1 (online data supplement) shows that total protein yield (mg per g sputum) consistently decreased for each CF participant following antibiotic treatment of exacerbation and that follow-up samples collected during clinical stability more closely resembled pre-treatment samples with respect to yield. Individual proteins identified as correlating negatively with FEV_1 % predicted (n=17) or consistently at high abundance in CF (n=6) were analysed to compare the relative abundance (sample:reference) of matched CF pre-treatment and <24h post-treatment samples. As shown in Table 2, paired samples T-tests (including Benjamini-Hochberg multiple comparisons adjustment of false discovery rate adjustment to <0.05) identified 9 proteins with decreased abundance (p<0.05) following treatment of pulmonary exacerbation. Figure 1 shows that for 7 of the 12 individuals with CF, samples collected during clinical stability (labelled ".3") clustered more closely with the matched pre-treatment sample (".1") than the post-treatment sample (".2").

ELISA validation of candidate biomarkers

As high abundance proteins (Figure 4) have greater potential for directly impacting the lung environment, we chose to validate myeloperoxidase and lactotransferrin as candidate biomarkers. Quantification of our existing samples by commercial ELISA confirmed their consistently higher abundance in CF compared to healthy control sputum cells (p<0.05, Mann Whitney U test, Figure E2 in the online supplement). Additionally, this demonstrated that commercial ELISAs are a practical method for follow-up validation of candidate biomarkers in a larger independent cohort.

DISCUSSION

In order to further elucidate the mechanisms of CF-associated tissue damage in adults with chronic respiratory *P. aeruginosa* infection, this study has for the first time specifically investigated the global proteome and functional activity of the combined population of cells present in CF sputum. As such, it uniquely studies the protein abundance of cells *in vivo*, thereby taking into account both their natural frequency in the sputum cell population and the effects of intercellular interactions. Thus, it delivers a high degree of clinical relevance to our findings. Moreover, to accurately characterise a consistent CF phenotype for adults chronically infected with *P. aeruginosa* and avoid bias due to patient individuality, we considered only those proteins detected in all samples within the CF and healthy control cohorts. Thus, our study is more stringent than previously published proteomic studies that either included proteins detected at high frequency in at least one sample [18] or used pooled samples [14-16].

In agreement with previous investigations of CF sputum or BAL [14, 15, 17, 18], we found the CF proteome to be largely distinct from that of the control cohort. IPA predicted that Canonical Pathways which affect or involve actin cytoskeleton rearrangement directed by the Rho family small GTPases, including RhoA, Rac1 and Cdc42 were over-represented in the CF samples. Thus, our *in vivo* data confirms previous *in vitro* observations linking CFTR deficiency with increased RhoA expression [27, 28]. This may be significant for CF lung health as impaired efferocytosis (phagocytosis of apoptotic cells) [29] and reduced expression of iNOS [28] resulting from CFTR deficiency have been reported as mediated by RhoA up-regulation.

Similarly to previous proteomics studies of cell-free CF sputum/BAL, GO and IPA Bio Function analyses identified up-regulation of the immune and inflammatory response [14-18]. In particular, we found increased immune cell movement and activation, and generation of reactive oxygen species. These observations fit with existing knowledge about the CF phenotype which is typified by chronic inflammation and neutrophilia. Interestingly, an IPA predicted network identified lipopolysaccharide as a likely trigger of inflammation. Cell death, growth and proliferation were also predicted as dysregulated in CF, in such a way that both death and proliferation of immune cells were decreased, while both were increased in non-immune cells. These observations further support the growing understanding of the CF phenotype. Tissue damage and cell death is rampant, while excessive proliferation of respiratory epithelial cells leading to squamous cell metaplasia and dysplasia is common in CF [30, 31]. Delayed neutrophil apoptosis and clearance is thought to contribute to chronic inflammation [2] and our study detected CF-associated differential abundance of matrix metalloproteinase 8, annexin I and nicotinamide phosphoribosyltransferase, which have been implicated in delayed apoptosis and clearing of inflammatory and immune cells *in vitro* [32] and *in vivo* [33, 34].

Compositional breakdown analysis of sample proteomes found S100-A9, the histone H2B family, lactotransferrin, histone H4, cathepsin G and myeloperoxidase to be consistently present in high abundance in CF and to each be significantly more abundant in the CF cohort than the control cohort. Importantly, all six proteins are known constituents of neutrophil extracellular trap (NET) structures, the release of which by neutrophils exerts antimicrobial activity through the degradation of virulence factors, bacterial killing and antifungal activity [35-37].

As our study focused on sputum cellular proteins we did not look specifically for NET structures, which are extracellular, and therefore do not know if they were present. However, NETs have been found previously in the CF airways fluid and the abundance of free DNA that is characteristic of NETs correlated with obstructive lung function [38]. As well as killing microbes, NETs have been found to directly induce human alveolar epithelial and endothelial cell death, with the predominant cytotoxic effect being attributed to histones and myeloperoxidase [39].

Individually, these six high abundance proteins may also have significance for CF lung pathology and several have previously been reported as up-regulated in CF sputum and/or BALF [14, 15, 18]. Importantly, differential abundance of myeloperoxidase in CF sputum fluid has been correlated with FEV_1 % predicted [17], and myeloperoxidase and calgranulin B have been identified in sputum fluid as CF-associated autoantigens [16]. Cathepsin G has been implicated in degradation of surfactant protein A, so reducing the innate pulmonary antimicrobial defence [40], while S100-A9 has been shown to induce MUC5AC production in human bronchial epithelial cells cultured *in vitro*, suggesting a potential contribution to pathological mucin hyperproduction *in vivo* [41]. In a study investigating trauma-induced lung injury, Abrams et al. reported that histones induced NET formation and myeloperoxidase release *in vitro*, while histone infusion in mouse models resulted in lung structural damage and neutrophil congestion [42]. Additionally, extracellular histones H4 and H3.3 have been implicated in asthma-related inflammation and remodeling, and in chronic obstructive pulmonary disease (COPD), respectively [43, 44]. Consequently, our consistent detection of these proteins in such high abundance in the sputum cellular fraction identifies a potential risk to lung health,

either through their active secretion and NET formation or following passive release into the sputum fluid via cell death.

A number of proteins detected consistently in CF sputum correlated negatively with lung function measured as FEV_1 % predicted. Their biological/clinical relevance and potential for causing injury are discussed more fully in the online supplement. Longitudinal analysis of patient-matched CF samples showed differential abundances for 10 out of these 17 proteins in the CF cohort following intravenous antibiotic treatment of a pulmonary exacerbation. Interestingly, all the changes in abundance resulted in levels closer to those in the control cohort. Previously, Sloane et al [17] observed that for 4/13 adults with CF, their cell-free sputum proteome profiles moved closer to those of healthy controls following antibiotic treatment of an exacerbation. However, the latter observations were made by comparison of 2-D gel images rather than implicating and identifying particular proteins. We also found that, although sputum total protein yield decreased in the CF cohort by completion of antibiotic treatment, it had reverted to pre-treatment levels by the clinically stable follow-up time point. This suggests that proteomic changes associated with antibiotic treatment may be temporary.

Overall, the proteins detected as negatively correlating with FEV_1 suggest that the degree of inflammation, neutrophil influx, protease activity and airway remodeling may all contribute to defining lung function. Alpha-1-antitrypsin and olfactomedin-4 were identified as independent factors negatively correlating with FEV_1 although this does not demonstrate causation of harm. Acute inflammation-associated protein alpha-1-antitrypsin has previously been identified as a biomarker of pulmonary disease in CF which can be predictive of response to treatment and prognosis [45]. As an inhibitor of neutrophil serine proteases, which contribute significantly to lung

destruction in CF, its negative correlation with lung function might seem surprising. However, previous studies observed that plasma deficiencies during inflammation were associated with less severe lung disease in CF [46]. No association between olfactomedin-4 and CF lung health has previously been reported. However, olfactomedin-4 has been shown to down-regulate proinflammatory responses to bacterial infection and reduce bacterial killing mediated by granule serine proteinases [47, 48], which could have the potential to negatively affect lung health.

There are a number of limitations to our study. As is typical for untargeted exploratory proteomic studies, we used highly sensitive shotgun technologies to obtain quantitative, high proteome coverage analysis of sputum cells as a novel target. The trade-off for such in-depth analysis is a relatively low sample size. Given the sample numbers practicable, we maximised power by restricting the study to adults with CF chronically infected with *P. aeruginosa*, who experienced pulmonary exacerbation requiring hospitalization for treatment with intravenous antibiotics. While CF sputum proteome responses to other microbial infections may vary, we considered this cohort most informative as 57% of the adult CF population is chronically infected with *P. aeruginosa* [49]. As this study investigated only adults the findings cannot be extended to children or individuals not chronically infected with *P aeruginosa*. Although proteins reported as potentially damaging were detected as consistently abundant or correlating negatively with lung function in our study, this is not proof of causation of lung injury. Additionally, the limited sample size requires that any potential significance as biomarkers either of lung function or disease severity warrants validation in a larger independent cohort. As an intial stage in this validation process we used commercial ELISAs to confirm findings in our existing cohort for myeloperoxidase and lactotransferrin. Both proteins were considered as

candidate biomarkers due to their consistently high abundance in CF sputum, increased abundance in the CF compared to the control cohort and previous associations in the literature with potential for lung damage. Thus, we demonstrated that commercial ELISAs are a practical method for follow-up validation in an independent cohort. However, while cohort size and characteristics limit the scope of our study conclusions, the findings are of exploratory interest and may have future clinical value within a defined commonly encountered CF cohort. Moreover, the investigation of all types of cells in the proportions occurring naturally in sputum is novel and the approach adopted in our study could be informative for other respiratory conditions.

In summary, within the cohort studied, i.e. adults with chronic *P. aeruginosa* infection, we have detected a distinct signature for the CF sputum cellular proteome, and have used pathway and functional analyses to predict its potential clinical impact. In particular, our data highlight the prevalence and predominance of proteins that can be associated with NET structures and have been reported as linked to tissue damage. Additionally, we identified a number of proteins that correlate negatively with lung function, many of which have previously been associated with lung injury. Overall, these findings expand the current understanding of the mechanisms behind *P. aeruginosa* associated CF lung disease in adults and identify sputum cell proteins as candidate biomarkers for disease status, prognostic indicators or as stratification parameters for treatment prescription.

ACKNOWLEDGMENTS

Sources of support: SHP was funded by the Cystic Fibrosis Trust Project Grant PJ545. DP, SP and the proteomics work were part supported by the NIH/NCI Cancer Center Support Grant 5P30CA045508.

REFERENCES

1 Boucher RC. Evidence for airway surface dehydration as the initiating event in CF airway disease. J Intern Med 2007;261:5-16.

2 Cohen TS, Prince A. Cystic fibrosis: a mucosal immunodeficiency syndrome. Nat Med 2012;18:509-519.

3 Cox MJ, Allgaier M, Taylor B, Baek MS, Huang YJ, Daly RA, Karaoz U, Andersen GL, Brown R, Fujimura KE, Wu B, Tran D, Koff J, Kleinhenz ME, Nielson D, Brodie EL, Lynch SV. Airway microbiota and pathogen abundance in age-stratified cystic fibrosis patients. PLoS One 2010;5:e11044.

4 Guss AM, Roeselers G, Newton IL, Young CR, Klepac-Ceraj V, Lory S, Cavanaugh CM. Phylogenetic and metabolic diversity of bacteria associated with cystic fibrosis. ISME J 2011;5:20-29.

5 Delhaes L, Monchy S, Fréalle E, Hubans C, Salleron J, Leroy S, Prevotat A, Wallet F, Wallaert B, Dei-Cas E, Sime-Ngando T, Chabé M, Viscogliosi E. The airway microbiota in cystic fibrosis: a complex fungal and bacterial community - implications for therapeutic management. PLoS One. 2012;7:e36313.

6 Bensalem N, Ventura AP, Vallée B, Lipecka J, Tondelier D, Davezac N, Dos Santos A, Perretti M, Fajac A, Sermet-Gaudelus I, Renouil M, Lesure JF, Halgand F, Laprévote O, Edelman A. Down-regulation of the anti-inflammatory protein annexin A1 in cystic fibrosis knock-out mice and patients. Mol Cell Proteomics 2005;4:1591-1601.

7 Roxo-Rosa M, da Costa G, Luider TM, Scholte BJ, Coelho AV, Amaral MD, Penque D. Proteomic analysis of nasal cells from cystic fibrosis patients and non-cystic fibrosis control individuals: search for novel biomarkers of cystic fibrosis lung disease. Proteomics 2006;6:2314-2325.

8 Carvalho-Oliveira I, Charro N, Aarbiou J, Buijs-Offerman RM, Wilke M, Schettgen T, Kraus T, Titulaer MK, Burgers P, Luider TM, Penque D, Scholte BJ. Proteomic analysis of naphthalene-induced airway epithelial injury and repair in a cystic fibrosis mouse model. J Proteome Res 2009;8:3606-3616.

9 Gomes-Alves P, Imrie M, Gray RD, Nogueira P, Ciordia S, Pacheco P, Azevedo P, Lopes C, de Almeida AB, Guardiano M, Porteous DJ, Albar JP, Boyd AC, Penque D. SELDI-TOF biomarker signatures for cystic fibrosis, asthma and chronic obstructive pulmonary disease. Clin Biochem 2010;43:168-177.

10 Ciavardelli D, D'Orazio M, Pieroni L, Consalvo A, Rossi C, Sacchetta P, Di Ilio C, Battistoni A, Urbani A. Proteomic and ionomic profiling reveals significant alterations of protein expression and calcium homeostasis in cystic fibrosis cells. Mol Biosyst 2013;9:1117-1126.

11 Ruaniyar N, Gupta V, Balch WE, Yates JR 3rd. Quantitative proteomic profiling reveals differentially regulated protein in cystic fibrosis cells. J Proteome Res 2014;13:4668-4675.

12 Lavallée-Adam M, Rauniyar N, McClatchy DB, Yates JR 3rd. PSEA-Quant: a protein set enrichment analysis on label-free and label-based protein quantification data. J Proteome Res 2014;13:5496-5509.

13 Jeanson L, Guerrera IC, Papon JF, Chhuon C, Zadigue P, Prulière-Escabasse V, Amselem S, Escudier E, Coste A, Edelman A. Proteomic analysis of nasal epithelial cells from cystic fibrosis patients. PLoS One 2014;9:e108671.

14 Gray RD, MacGregor G, Noble D, Imrie M, Dewar M, Boyd AC, Innes JA, Porteous DJ, Greening AP. Sputum proteomics in inflammatory and suppurative respiratory diseases. Am J Respir Crit Care Med 2008;178:444-452.

15 MacGregor G, Gray RD, Hilliard TN, Imrie M, Boyd AC, Alton EW, Bush A, Davies JC, Innes JA, Porteous DJ, Greening AP. Biomarkers for cystic fibrosis lung disease: Application of SELDI-TOF mass spectrometry to BAL fluid. J Cyst Fibros 2008;7:352-358.

16 Pedersen SK, Sloane AJ, Prasad SS, Sebastian LT, Lindner RA, Hsu M, Robinson M, Bye PT, Weinberger RP, Harry JL. An immunoproteomic approach for identification of clinical biomarkers for monitoring disease. Mol Cell Proteomics 2005;4:1052-1060.

17 Sloane AJ, Lindner RA, Prasad SS, Sebastian LT, Pedersen SK, Robinson M, Bye PT, Nielson DW, Harry JL. Proteomic analysis of sputum from adults and children with cystic fibrosis and from control subjects. Am J Respir Crit Care Med 2005;172:1416-1426.

18 Gharib S, Vaisar T, Aitken M, Park DR, Heinecke JW, Fu X. Mapping the lung proteome in cystic fibrosis. J Proteome Res 2009;8:3020-3029.

19 Peters-Hall JR, Brown KJ, Pillai DK, Tomney A, Garvin LM, Wu X, Rose MC. Quantitative proteomics reveals an altered cystic fibrosis in vitro bronchial epithelial secretome. Am J Respir Cell Mol Biol 2015;53:22-32.

20 Pattison SH, Johnston E, Gibson DS, Peacock S, Pappin DJ, Elborn JS Proteomics analysis of the chronically infected CF airways [abstract]. J Cyst Fibros 2012;11:WS11.10.

21 Fuchs HJ, Borowitz DS, Christiansen DH, Morris EM, Nash ML, Ramsey BW, Rosenstein BJ, Smith AL, Wohl ME. Effect of aerosolized recombinant human DNase on exacerbations of respiratory symptoms and on pulmonary function in patients with cystic fibrosis. N Engl J Med 1994;331:637-642.

22 Scuoppo C, Miething C, Lindqvist L, Reyes J, Ruse C, Appelmann I, Yoon S, Krasnitz A, Teruya-Feldstein J, Pappin D, Pelletier J, Lowe SW. A tumour suppressor network relying on the polyamine-hypusine axis. Nature 2012;487:244-248.

23 Vizcaíno JA, Deutsch EW, Wang R, Csordas A, Reisinger F, Ríos D, Dianes JA, Sun Z, Farrah T, Bandeira N, Binz PA, Xenarios I, Eisenacher M, Mayer G, Gatto L, Campos A, Chalkley RJ, Kraus HJ, Albar JP, Martinez-Bartolomé S, Apweiler R, Omenn GS, Martens L, Jones AR, Hermjakob H. ProteomeXchange provides globally coordinated proteomics data submission and dissemination. Nat Biotechnol 2014;32:223-226.

24 Caraux G, Pinloche S. PermutMatrix: a graphical environment to arrange gene expression profiles in optimal linear order. Bioinformatics 2005;21:1280-1281.

25 Ishihama Y, Oda Y, Tabata T, Sato T, Nagasu T, Rappsilber J, Mann M. Exponentially modified Protein Abundance Index (emPAI) for estimation of absolute protein amount in proteomics by the number of sequenced peptides per protein. Mol Cell Proteomics 2005;4:1265-1272.

26 Uhlén M, Björling E, Agaton C, Szigyarto CA, Amini B, Andersen E, Andersson AC, Angelidou P, Asplund A, Asplund C, Berglund L, Bergström K, Brumer H, Cerjan D, Ekström M, Elobeid A, Eriksson C, Fagerberg L, Falk R, Fall J, Forsberg M, Björklund MG, Gumbel K, Halimi A, Hallin I, Hamsten C, Hansson M, Hedhammar M, Hercules G, Kampf C, Larsson K, Lindskog M, Lodewyckx W, Lund J, Lundeberg J, Magnusson K, Malm E, Nilsson P, Odling J, Oksvold P, Olsson I, Oster E, Ottosson J, Paavilainen L, Persson A, Rimini R, Rockberg J, Runeson M, Sivertsson A, Sköllermo A, Steen J, Stenvall M, Sterky F, Strömberg S, Sundberg M, Tegel H, Tourle S, Wahlund E, Waldén A, Wan J, Wernérus H, Westberg J, Wester K, Wrethagen U, Xu LL, Hober S, Pontén F. A human protein atlas for normal and cancer tissues based on antibody proteomics. Mol Cell Proteomics 2005;4:1920-1932.

27 Perez A, Issler AC, Cotton CU, Kelley TJ, Verkman AS, Davis PB. CFTR inhibition mimics the cystic fibrosis inflammatory profile. Am J Physiol Lung Cell Mol Physiol 2007;292:L383-395.

28 Kreiselmeier NE, Kraynack NC, Corey DA, Kelley TJ. Statin-mediated correction of STAT1 signaling and inducible nitric oxide synthase expression in cystic fibrosis epithelial cells. Am J Physiol Lung Cell Mol Physiol 2003;285:L1286-1295.

29 Vandivier RW, Richens TR, Horstmann SA, deCathelineau AM, Ghosh M, Reynolds SD, Xiao YQ, Riches DW, Plumb J, Vachon E, Downey GP, Henson PM. Dysfunctional cystic fibrosis transmembrane conductance regulator inhibits phagocytosis of apoptotic cells with proinflammatory consequences. Am J Physiol Lung Cell Mol Physiol 2009;297:677-686.

30 Voynow JA, Fischer BM, Roberts BC, Proia AD. Basal-like cells constitute the proliferating cell population in cystic fibrosis airways. Am J Respir Crit Care Med 2005;172:1013-1018.

31 Hajj R, Lesimple P, Nawrocki-Raby B, Birembaut P, Puchelle E, Coraux C. Human airway surface epithelial regeneration is delayed and abnormal in cystic fibrosis. J Pathol 2007;211:340-350.

32 Scannell M, Flanagan MB, deStefani A, Wynne KJ, Cagney G, Godson C, Maderna P. Annexin-1 and peptide derivatives are released by apoptotic cells and stimulate phagocytosis of apoptotic neutrophils by macrophages. J Immunol 2007;178:4595-4605.

33 Jia SH, Li Y, Parodo J, Kapus A, Fan L, Rotstein OD, Marshall JC. Pre–B cell colony–enhancing factor inhibits neutrophil apoptosis in experimental inflammation and clinical sepsis. J Clin Invest 2004;113:1318-1327.

34 Gueders MM, Balbin M, Rocks N, Foidart JM, Gosset P, Louis R, Shapiro S, Lopez-Otin C, Noël A, Cataldo DD. Matrix metalloproteinase-8 deficiency promotes granulocytic allergen-induced airway inflammation. J Immunol 2005;175:2589-2597.

35 Brinkmann V, Reichard U, Goosmann C, Fauler B, Uhlemann Y, Weiss DS, Weinrauch Y, Zychlinsky A. Neutrophil extracellular traps kill bacteria. Science 2004;303:1532-1535.

36 Urban CF, Ermert D, Schmid M, Abu-Abed U, Goosmann C, Nacken W, Brinkmann V, Jungblut PR, Zychlinsky A. Neutrophil extracellular traps contain calprotectin, a cytosolic protein complex involved in host defense against Candida albicans. PLoS Pathog 2009;5:e1000639.

37 Bianchi M, Niemiec MJ, Siler U, Urban CF, Reichenbach J. Restoration of anti-Aspergillus defense by neutrophil extracellular traps in human chronic granulomatous disease after gene therapy is calprotectin-dependent. J Allergy Clin Immunol 2011;127:1243-1252.

38 Marcos V, Zhou Z, Yildirim AO, Bohla A, Hector A, Vitkov L, Krautgartner WD, Stoiber W, Griese M, Eickelberg O, Mall MA, Hartl D. Free DNA in cystic fibrosis airway fluids correlates with airflow obstruction. Mediators Inflamm 2015; 2015:408935.

39 Saffarzadeh M, Juenemann C, Queisser MA, Lochnit G, Barreto G, Galuska SP, Lohmeyer J, Preissner KT. Neutrophil extracellular traps directly induce epithelial and endothelial cell death: a predominant role of histones. PLoS One 2012;7:e32366.

40 Rubio F, Cooley J, Accurso FJ, Remold-O'Donnell E. Linkage of neutrophil serine proteases and decreased surfactant protein-A (SP-A) levels in inflammatory lung disease. Thorax 2004;59:318-323.

41 Kang JH, Hwang SM, Chung IY. S100A8, S100A9 and S100A12 activate airway epithelial cells to produce MUC5AC via extracellular signal-regulated kinase and nuclear factor-κB pathways. Immunology 2015;144:79-90.

42 Abrams ST, Zhang N, Manson J, Liu T, Dart C, Baluwa F, Wang SS, Brohi K, Kipar A, Yu W, Wang G, Toh CH. Circulating histones are mediators of trauma-associated lung injury. Am J Respir Crit Care Med 2013;187:160-169.

43 Calvo FQ, Fillet M, de Seny D, Meuwis MA, Maree R, Crahay C, Paulissen G, Rocks N, Gueders M, Wehenkel L, Merville MP, Louis R, Foidart JM, Noël A, Cataldo D. Biomarker discovery in asthma-related inflammation and remodeling. Proteomics 2009;9:2163-2170.

44 Barrero CA, Perez-Leal O, Aksoy M, Moncada C, Ji R, Lopez Y, Mallilankaraman K, Madesh M, Criner GJ, Kelsen SG, Merali S. Histone 3.3 participates in a self-sustaining cascade of apoptosis that contributes to the progression of chronic obstructive pulmonary disease. Am J Respir Crit Care Med 2013;188:673-683.

45 Sagel SD, Thompson V, Chmiel JF, Montgomery GS, Nasr SZ, Perkett E, Saavedra MT, Slovis B, Anthony MM, Emmett P, Heltshe SL. Effect of treatment of cystic fibrosis pulmonary exacerbations on systemic inflammation. Ann Am Thorac Soc 2015;12:708-717.

46 Mahadeva R, Westerbeek R, Perry D, Lovegrove JU, Whitehouse DB, Carroll NR, Ross-Russell RI, Webb AK, Bilton D, Lomas DA. Alpha1-antitrypsin deficiency alleles and the Taq-I G-->A allele in cystic fibrosis lung disease. Eur Respir J 1998;11:873-879.

47 Liu W, Yan M, Liu Y, Wang R, Li C, Deng C, Singh A, Coleman WG Jr, Rodgers GP. Olfactomedin 4 down-regulates innate immunity against Helicobacter pylori infection. Proc Natl Acad Sci USA 2010;107:11056-11061.

48 Liu W, Yan M, Liu Y, McLeish KR, Coleman WG Jr, Rodgers GP. Olfactomedin 4 inhibits cathepsin C-mediated protease activities, thereby modulating neutrophil killing of Staphylococcus aureus and Escherichia coli in mice. J Immunol 2012;189:2460-2467.

49 Cystic Fibrosis Foundation Patient Registry: 2015 Annual Data Report to the Center Directors. Bethesda, Maryland: Cystic Fibrosis Foundation; 2016.

Figure 1: Distribution and relative abundance of human proteins detected in sputum samples collected from CF and healthy control cohorts. A. shows that 119 proteins were found in all 48 samples, i.e. both CF and control, while a further 21 proteins were found exclusively in all CF samples and 17 proteins exclusively in all control samples. B. shows hierarchical clustering (using Euclidean distance dissimilarity and McQuitty's linkage criteria) of individual sputum samples according to the relative abundance levels for all 119 consistently detected proteins. CF and control samples largely cluster separately with the only exceptions being 2 CF samples (marked with black arrows) collected after antibiotic treatment for pulmonary exacerbation. The scale is a $\log_2$ ratio of relative abundance.

Figure 2: Differential activity of Canonical Pathways in consensus CF and control proteomes. Based on the set of proteins detected in all samples within a cohort, Ingenuity Pathway Analysis software applied Fisher's Exact Test to predict the Canonical Pathways likely to be active within that cohort ($p<0.05$). The top 5 pathways predicted in the control cohort concern general metabolism, clatherin-mediated endocytosis signaling and mitochondrial dysfunction and these were detected with approximately equal probabilities in both cohorts. Pathways where there is a >5-fold increase in *p*-value in the CF proteome compared to the control proteome, indicating likely up-regulation in CF are involved in immune functions including neutrophil recruitment, rearrangement of the actin cytoskeleton, phagocytosis and T cell signaling.

Figure 3: Ingenuity Pathway Analysis predicted network for proteins detected in CF involved in Inflammatory Response and Cell Movement Bio Functions. The intensity of red or green shading indicates the degree of increased or decreased abundance in CF relative to healthy controls, respectively. A red outline with no internal shading indicates that the protein was detected exclusively in the CF cohort. The network identifies lipopolysaccharide (LPS) as a likely inflammatory trigger.

Figure 4: Percentage abundance estimates (% emPAI median ± IQR) for proteins comprising >1% of total human proteins in >80% of CF samples. Human Protein Atlas (26) annotations for protein expression in terms of tissue location and localised abundance are recorded. H, M and L indicate high, medium and low protein expression. All proteins are either highly or moderately expressed in bone marrow hematopoietic cells, 50% of which are myelopoietic cells, and all except cathepsin G are moderately or lowly expressed in lung macrophages.

Proteomic profile of cystic fibrosis sputum cells in adults chronically infected with *Pseudomonas aeruginosa*

Sally H. Pattison, David S. Gibson, Elinor Johnston, Samantha Peacock, Keith Rivera, Michael M. Tunney, Darryl J. Pappin, and J. Stuart Elborn

Supplementary Data

METHODS

Extraction of sputum cellular proteins

All work was carried out at 4°C and using prechilled buffers. Approximately 1g of mucus plug was separated from each sputum sample and mixed with 9 ml sterile phosphate buffered saline pH 7.4 (PBS) containing Roche Complete Mini protease inhibitor cocktail (PIC). The mixture was homogenised thoroughly by repeat syringing (no needle) and cells pelleted by centrifugation for 10 min at 4,000 x g. The cells were washed twice more in 10 ml PBS+PIC and strained through a 70 µm nylon cell strainer (BD Biosciences) to remove any remaining cell clumps. The cells were pelleted once more to remove all supernatant before protein extraction. Cell pellets were resuspended in 0.2 - 3.0 ml lysis buffer (10 mM HEPES, 0.5 mM EDTA, 1% v/v Nonidet P 40, 4% w/v SDS, pH 8.0) depending on pellet size and Sigma protease inhibitor cocktail added at 10 µl per 1 ml as directed. After thorough homogenisation by vortexing and repeat passage through an 18-gauge needle, the mixtures were subjected to 5 freeze-thaw cycles and water bath sonication to ensure lysis. Insoluble cell debris was removed by centrifugation at 14,000 x g and the soluble protein supernatants were stored at -80°C.

Protein digestion and iTRAQ labelling

Protein digestion and iTRAQ labelling were performed as described previously (1). Protein concentrations were determined by BCA Protein Assay (Pierce). ProteaseMax (Promega) made up in 50 mM triethylammonium bicarbonate (TEAB) was added to a final concentration of 0.1% to 70 µg of each protein sample. The samples were reduced with 5 mM tris-(2-carboxyethyl) phosphine at 55°C for 20 min and alkylated with 10 mM methyl methanethiosulphonate at room temperature for 20 min. Following methanol/chloroform precipitation for 60 min at 4°C, protein pellets

were resuspended in 50 mM TEAB, 0.1% Proteasemax and trypsin digested overnight at 37°C. The tryptic peptides were dried to <5 µl by vacuum centrifugation and diluted with 500 mM TEAB to a final volume of 30 µl, before acidification with trifluoroacetic acid and labelling with 8plex iTRAQ Reagents (AB Sciex) according to kit instructions.

MudPIT

Six test samples and a reference sample were mixed for simultaneous analysis by 2D liquid chromatography MudPIT, using a two-dimensional Vented Column Setup with a Proxeon nano-flow high-performance liquid pump as described previously (1). Eight MudPIT runs were performed in total and the samples combined per run are recorded in Table E6. Each salt step was eluted at 3.5 µl/min on an analytical column (100 µm internal diameter packed with 12 cm of 5 µm Aqua C18) on line with an LTQ Orbitrap Velos (Thermo Electron). Data were acquired in profile mode using the following parameters: for full-scan Fourier transform mass spectrometry, resolution = 60,000, *m/z* = 380–1,700 and the 10 most intense ions were fragmented with higher-collision dissociation at a normalised collision energy of 40% and an activation time of 0.1. Minimum threshold signal was at 5,000 and isolation width at 1.2. Dynamic exclusion settings were repeat count 1, repeat duration of 30, exclusion list size 500, exclusion duration 60 and exclusion mass width 10 p.p.m.

Protein identification

Peaklists were generated by Mascot Distiller (Matrix science; version 2.3). Protein identification was performed by the Mascot search engine version 2.3 against the UniProt human database 2016 (92,179 sequences) plus NCBI non-redundant database (2016), taxonomy *P. aeruginosa* (634,627 sequences) with a precursor mass tolerance of 20 p.p.m., with iTRAQ8plex (N-terminal) and iTRAQ8plex (lysine)

as fixed modifications and oxidation (methionine) and deamidated (asparagine & lysine) as variable modifications. Trypsin was set as the cleavage enzyme and one missed cleavage was allowed. A mass tolerance of 0.3 Da was set for fragment ions. Mascot Percolator was used to re-score the results to false discovery rates of 0.74-1.16% for protein identification. Protein identifications with a MOWSE score less than 65 were rejected. Protein-level iTRAQ ratios were calculated as intensity weighted, using only peptides with expectation values less than 0.05. Global ratio normalization was performed using intensity summation, with no outlier rejection.

Enzyme-linked Immunosorbent assay (ELISA)

In order to validate myeloperoxidase and lactotransferrin as high abundance proteins in the CF cohort compared to control cohort and to demonstrate that commercial ELISAs are a practical method for validation of candidate biomarkers, both proteins were quantified in each sample by ELISA (ab119605 and ab108882 respectively, Abcam, Cambridge, UK) according to the manufacturer's instructions. Figure E2 shows the median±IQR % abundance for each candidate biomarker along with the corresponding data for the control cohort.

Western immunoblot

Western immunoblot was used to further confirm the mass spectrometry identification and quantification data (Figure E3). Patient sample protein concentrations were assessed by BCA assay (Thermo Scientific Inc., Rockford, USA). 25 µg of protein was loaded per lane on 4-20% TGX™ gels and transferred to 0.45 µm PVDF membrane (Millipore Corproration, Massachusetts, USA) with a Criterion™ mini gel tank and blot module, associated buffers and power supply, according to the manufacturer's guidelines (Bio-rad Laboratories Inc., Hercules, USA). Nonspecific binding was blocked overnight at 4 °C using 5% non-fat milk in PBS with 0.05% Tween 20 (PBST) (Sigma-Aldrich Company Ltd., Gillingham, UK). Primary antibodies against azurocidin, heatshock protein beta 1 and cytokeratin (Abcam Ltd, Cambridge, UK) were incubated with blotted membranes at a diluted in PBST, for 1 h at room temperature. Membranes were washed 3 times (5 min each) with PBST and incubated in secondary anti-IgG horseradish peroxidase diluted in PBST (Promega Inc., Madison, USA), for 1 h at room temperature. After a further 6 washes, bands were visualized using Supersignal West Pico chemiluminescent substrate (Thermo Scientific Inc., Rockford, USA), imaged with a Autochemi CCD camera and analysed with Labworks software version 4.0.0.8 (UVP Ltd., Cambridge, UK).

DISCUSSION

A number of proteins detected consistently in CF sputum correlated negatively with lung function measured as FEV_1 % predicted. Their biological/clinical relevance and potential for causing injury are discussed below. Interestingly, 9 out of these 15 proteins also showed differential abundance in the CF cohort following intravenous antibiotic treatment of a pulmonary exacerbation, with the changes in abundance bringing them closer to levels in the control cohort.

Acute inflammation-associated proteins alpha-1-antitrypsin, haptoglobin, fibrinogen and alpha-2-macroglobulin have previously been identified as serum biomarkers of pulmonary disease including CF and COPD and can be predictive of response to treatment and prognosis (2-6). The presence of these proteins in the lungs has traditionally been thought to be due to plasma leakage. However, haptoglobin and fibrinogen may also be synthesized in the lungs in response to inflammation and injury, and may exert roles in immunity and repair mechanisms (7, 8). Alpha-1-antitrypsin is an inhibitor of neutrophil serine proteases, including elastase, which contribute significantly to lung destruction in CF, and therefore its negative correlation with lung function might seem surprising. However, previous studies observed that plasma deficiencies during inflammation were associated with less severe lung disease in CF (9, 10). Fibrinogen has been shown to increase mucin production in airway epithelial cells (11), which could contribute to the CF hypersecretory phenotype.

Three proteins associated with the migration of neutrophils, lymphocytes and alveolar monocytes – integrin beta 2, vanin-2 and serine/threonine-protein kinase 10 (12-14) – correlated negatively with lung function. This correlation could be due to either the negative impact of these cells as part of an overactive immune system or

poor lung health stimulating an increased influx of immune cells. In accordance with the predicted neutrophil influx in CF, four granule proteins – neutrophil elastase, olfactomedin-4, CD63 antigen and maltase-glucoamylase – were also found to correlate negatively with lung function. Neutrophil elastase exerts antimicrobial activity and therefore might be expected to promote lung health. However, CF-associated accumulation of neutrophil elastase can be damaging to lung tissue (15-17) and sputum measurements have previously been shown to predict the rate of FEV_1 decline in children with CF (18), while higher levels of neutrophil elastase following antibiotic treatment of an exacerbation have been associated with an increased risk of subsequent exacerbation (19). Additionally, serum levels of elastase/alpha-1-antitrypsin complex have been found to predict response to treatment of chronic *P. aeruginosa* colonisation (3). In contrast, olfactomedin-4 has been shown to down-regulate proinflammatory responses to bacterial infection and reduce bacterial killing mediated by granule serine proteinases (20, 21), which would be expected to negatively affect lung health. Additionally, olfactomedin-4 has recently been identified as a novel target autoantigen for antineutrophil cytoplasmic antibodies (ANCA), although both cases reported were for people without CF (22).

Neprilysin cleaves inflammatory peptides and has been proposed as protective against hypoxia- and smoke-related pulmonary vascular remodeling in COPD (23, 24), while thioredoxin-dependent peroxidase reductase has been implicated as an important scavenger of reactive oxygen species during LPS-induced oxidative stress and gene knockout mice showed increased lung inflammation (25). Although both correlated negatively with lung function, there is no published data suggesting their effects can be harmful. Thus, it is likely that observed levels signify a response to inflammation rather than a cause. BASP1 is strongly expressed in ciliated respiratory

cells (26) and its detection in sputum cells is likely indicative of significant dislodgement of these cells and injury to the respiratory tract that would be expect to hinder mucus clearance. Additionally, increased BASP1 expression has been linked to apoptosis of human tubular epithelial cells and in diabetic nephropathy (27). Finally, levels of RhoA also correlated negatively with lung function. As already discussed, RhoA is involved in the regulation of many activities relating to inflammatory and immune response repair such as phagocytosis, cell migration and wound closure. In part its negative correlation with lung function may reflect the degree of injury, but also its promotion of neutrophil influx may incur a mechanism for tissue damage.

REFERENCES

1 Scuoppo C, Miething C, Lindqvist L, Reyes J, Ruse C, Appelmann I, Yoon S, Krasnitz A, Teruya-Feldstein J, Pappin D, *et al.* A tumour suppressor network relying on the polyamine-hypusine axis. Nature 2012;487:244-248.

2 Engström G, Segelstorm N, Ekberg-Aronsson M, Nilsson PM, Lindgärde F, Löfdahl C. Plasma markers of inflammation and incidence of hospitalisations for COPD: results from a population-based cohort study. Thorax 2009;64:211-215.

3 Valletta EA, Rigo A, Bonazzi L, Zanolla L, Mastella G. Modification of some markers of inflammation during treatment for acute respiratory exacerbation in cystic fibrosis. Acta Paediatr 1992;81:227-230.

4 Duvoix A, Dickens J, Haq I, Mannino D, Miller B, Tal-Singer R, Lomas DA. Blood fibrinogen as a biomarker of chronic obstructive pulmonary disease. Thorax 2013;68:670-676.

5 Jiang R, Burke GL, Enright PL, Newman AB, Margolis HG, Cushman M, Tracy RP, Wang Y, Kronmal RA, Barr RG. Inflammatory markers and longitudinal lung function decline in the elderly. Am J Epidemiol 2008;168:602-610.

6 Sagel SD, Thompson V, Chmiel JF, Montgomery GS, Nasr.S Z, Perkett E, Saavedra MT, Slovis B, Anthony MM, Emmett P, *et al.* Effect of treatment of cystic fibrosis pulmonary exacerbations on systemic inflammation. Ann Am Thorac Soc 2015;12:708-717.

7 Abdullah M, Kähler D, Vock C, Reiling N, Kugler C, Drömann D, Rupp J, Hauber HP, Fehrenbach H, Zabel P, *et al.* Pulmonary haptoglobin and CD163 are functional immunoregulatory elements in the human lung. Respiration 2012;83:61-73.

8 Guadiz G, Sporn LA, Goss RA, Lawrence SO, Marder VJ, Simpson-Haidaris PJ. Polarized secretion of fibrinogen by lung epithelial cells. Am J Respir Cell Mol Biol 1997;17:60-69.

9 Mahadeva R, Westerbeek R, Perry D, Lovegrove J, Whitehouse D, Carroll N, Ross-Russell R, Webb A, Bilton D, Lomas D. Alpha1-antitrypsin deficiency alleles and the Taq-I G-->A allele in cystic fibrosis lung disease. Eur Respir J 1998;11:873-879.

10 Mahadeva R, Stewart S, Bilton D, Lomas DA. Alpha-1 antitrypsin deficiency alleles and severe cystic fibrosis lung disease. Thorax 1998;53:1022-1024.

11 Kim S, Nadel JA. Fibrinogen binding to ICAM-1 promotes EGFR-dependent mucin production in human airway epithelial cells. Am J Physiol Lung Cell Mol Physiol 2009;297:L174-183.

12 Rosseau S, Selhorst J, Wiechmann K, Leissner K, Maus U, Mayer K, Grimminger F, Seeger W, Lohmeyer J. Monocyte migration through the alveolar epithelial barrier: adhesion molecule mechanisms and impact of chemokines. J Immunol 2000;164:427-435.

13 Suzuki K, Watanabe T, Sakurai S, Ohtake K, Kinoshita T, Araki A, Fujita T, Takei H, Takeda Y, Sato Y, *et al.* A novel glycosylphosphatidyl inositol-anchored protein on human leukocytes: a possible role for regulation of neutrophil adherence and migration. J Immunol 1999;162:4277-4284.

14 Belkina NV, Liu Y, Hao J, Karasuyama H, Shaw S. LOK is a major ERM kinase in resting lymphocytes and regulates cytoskeletal rearrangement through ERM phosphorylation. Proc Natl Acad Sci USA 2009;106:4707-4712.

15 Amitani R, Wilson R, Rutman A, Read R, Ward C, Burnett D, Stockley RA, Cole PJ. Effects of human neutrophil elastase and Pseudomonas aeruginosa proteinases on human respiratory epithelium. Am J Respir Cell Mol Biol 1991;4:26-32.

16 O'Connor CM, Gaffney K, Keane J, Southey A, Byrne N, O'Mahoney S, Fitzgerald MX. alpha 1-Proteinase inhibitor, elastase activity, and lung disease severity in cystic fibrosis. Am Rev Respir Dis 1993;148:1665-1670.

17 Gehrig S, Duerr J, Weitnauer M, Wagner CJ, Graeber SY, Schatterny J, Hirtz S, Belaaouaj A, Dalpke AH, Schultz C, *et al.* Lack of neutrophil elastase reduces inflammation, mucus hypersecretion, and emphysema, but not mucus obstruction, in mice with cystic fibrosis-like lung disease. Am J Respir Crit Care Med 2014;189:1082-1092.

18 Sagel SD, Wagner BD, Anthony MM, Emmett P, Zemanick ET. Sputum biomarkers of inflammation and lung function decline in children with cystic fibrosis. Am J Respir Crit Care Med 2012;186:857-865.

19 Waters VJ, Stanojevic S, Sonneveld N, Klingel M, Grasemann H, Yau YCW, Tullis E, Wilcox P, Freitag A, Chilvers M, *et al.* Factors associated with response to treatment of pulmonary exacerbations in cystic fibrosis patients [published online ahead of print February 14, 2015]. J Cyst Fibros ;doi: 10.1016/j.jcf.2015.01.007.

20 Liu W, Yan M, Liu Y, Wang R, Li C, Deng C, Singh A, Coleman WG, Rodgers GP. Olfactomedin 4 down-regulates innate immunity against Helicobacter pylori infection. Proc Natl Acad Sci USA 2010;107:11056-11061.

21 Liu W, Yan M, Liu Y, McLeish KR, Coleman WG, Rodgers GP. Olfactomedin 4 inhibits cathepsin C-mediated protease activities, thereby modulating neutrophil killing of Staphylococcus aureus and Escherichia coli in mice. J Immunol 2012;189:2460-2467.

22 Amirbeagi F, Thulin P, Pullerits R, Pedersen B, Andersson BA, Dahlgren C, Welin A, Bylund J. Olfactomedin-4 autoantibodies give unusual c-ANCA staining patterns with reactivity to a subpopulation of neutrophils. J Leukoc Biol 2015;97:181-189.

23 Wick MJ, Buesing EJ, Wehling CA, Loomis ZL, Cool CD, Zamora MR, Miller YE, Colgan SP, Hersh LB, Voelkel NF, *et al.* Decreased neprilysin and pulmonary vascular remodeling in chronic obstructive pulmonary disease. Am J Respir Crit Care Med 2011;183:330-340.

24 Hashimoto S, Amaya F, Oh-Hashi K, Kiuchi K, Hashimoto S. Expression of neutral endopeptidase activity during clinical and experimental acute lung injury. Respir Res 2010;11:164.

25 Li L, Shoji W, Takano H, Nishimura N, Aoki Y, Takahashi R, Goto S, Kaifu T, Takai T, Obinata M. Increased susceptibility of MER5 (peroxiredoxin III) knockout mice to LPS-induced oxidative stress. Biochem Biophys Res Commun 2007;355:715-721.

26 Uhlén M, Björling E, Agaton C, Szigyarto CA, Amini B, Andersen E, Andersson A, Angelidou P, Asplund A, Asplund C, *et al.* A human protein atlas for normal and cancer tissues based on antibody proteomics. Mol Cell Proteomics 2005;4:1920-1932.

27 Sanchez-Niño MD, Sanz AB, Lorz C, Gnirke A, Rastaldi MP, Nair V, Egido J, Ruiz-Ortega M, Kretzler M, Ortiz A. BASP1 promotes apoptosis in diabetic nephropathy. J Am Soc Nephrol 2010;21:610-621.

28 Bensalem N, Ventura AP, Vallée B, Lipecka J, Tondelier D, Davezac N, Santos AD, Perretti M, Fajac A, Sermet-Gaudelus I, *et al.* Down-regulation of the anti-inflammatory protein annexin A1 in cystic fibrosis knock-out mice and patients. Mol Cell Proteomics 2005;4:1591-1601.

29 Roxo-Rosa M, da Costa G, Luider T, Scholte B, Coelho A, Amaral M, Penque D. Proteomic analysis of nasal cells from cystic fibrosis patients and non-cystic fibrosis control individuals: search for novel biomarkers of cystic fibrosis lung disease. Proteomics 2006;6:2314-2325.

30 Carvalho-Oliveira I, Charro N, Aarbiou J, Buijs-Offerman R, Wilke M, Schettgen T, Kraus T, Titulaer M, Burgers P, Luider T, *et al.* Proteomic analysis of naphthalene-induced airway epithelial injury and repair in a cystic fibrosis mouse model. J Proteome Res 2009;8:3606-3616.

31 Gomes-Alves P, Imrie M, Gray RD, Nogueira P, Ciordia S, Pacheco P, Azevedo P, Lopes C, de Almeida AB, Guardiano M, *et al.* SELDI-TOF biomarker signatures for cystic fibrosis, asthma and chronic obstructive pulmonary disease. Clin Biochem 2010;43:168-177.

32 Ciavardelli D, D'Orazio M, Pieroni L, Consalvo A, Rossi C, Sacchetta P, Di Ilio C, Battistoni A, Urbani A. Proteomic and ionomic profiling reveals significant alterations of protein expression and calcium homeostasis in cystic fibrosis cells. Mol Biosyst 2013;9:1117-1126.

33 Ruaniyar N, Gupta V, Balch WE, Yates JR3. Quantitative proteomic profiling reveals differentially regulated protein in cystic fibrosis cells. J Proteome Res 2014;13:4668-4675.

34 Lavallée-Adam M, Rauniyar N, McClatchy DB, Yates JR3. PSEA-Quant: a protein set enrichment analysis on label-free and label-based protein quantification data. J Proteome Res 2014;13:5496-5509.

35 Jeanson L, Guerrera IC, Papon JF, Chhuon C, Zadigue P, Prulière-Escabasse V, Amselem S, Escudier E, Coste A, Edelman A. Proteomic analysis of nasal epithelial cells from cystic fibrosis patients. PLoS One 2014;9:e108671.

36 Peters-Hall JR, Brown KJ, Pillai DK, Tomney A, Garvin LM, Wu X, Rose MC. Quantitative proteomics reveals an altered cystic fibrosis in vitro bronchial epithelial secretome. Am J Respir Cell Mol Biol 2015;53:22-32.

37 Gray RD, MacGregor G, Noble D, Imrie M, Dewar M, Boyd AC, Innes JA, Porteous DJ, Greening AP. Sputum proteomics in inflammatory and suppurative respiratory diseases. Am J Respir Crit Care Med 2008;178:444-452.

38 Sloane AJ, Lindner RA, Prasad SS, Sebastian LT, Pedersen SK, Robinson M, Bye PT, Nielson DW, Harry JL. Proteomic analysis of sputum from adults and children with cystic fibrosis and from control subjects. Am J Respir Crit Care Med 2005;172:1416-1426.

39 Gharib S, Vaisar T, Aitken M, Park D, Heinecke J, Fu X. Mapping the lung proteome in cystic fibrosis. J Proteome Res 2009;8:3020-3029.

40 MacGregor G, Gray RD, Hilliard TN, Imrie M, Boyd AC, Alton EWFW, Bush A, Davies JC, Innes JA, Porteous DJ, *et al.* Biomarkers for cystic fibrosis lung disease: Application of SELDI-TOF mass spectrometry to BAL fluid. J Cyst Fibros 2008;7:352-358.

41 Pedersen SK, Sloane AJ, Prasad SS, Sebastian LT, Lindner RA, Hsu M, Robinson M, Bye PT, Weinberger RP, Harry JL. An immunoproteomic approach for identification of clinical biomarkers for monitoring disease. Mol Cell Proteomics 2005;4:1052-1060.

Sample type	CF phenotype/observation	Refs
human nasal epithelial cells	reduced expression of the anti-inflammatory protein annexin A1 in CF	(28)
human nasal epithelial cells	differential expression of proteins related to chronic inflammation, oxidative stress, cytoskeleton proteins and mitochondrial proteins	(29)
mouse airway epithelial cells	reduced expression of carbonyl reductase (prostaglandin metabolism) and a more pronounced down-regulation of two key enzymes of retinoic acid metabolism after injury	(30)
human nasal epithelial cells	multiple-biomarker signature patterns for CF	(31)
human bronchial epithelial cell line	differences in glycolytic and gluconeogenic pathways, unfolded protein responses and a signaling pathway involved in cellular calcium homeostasis	(32)
human bronchial epithelial cell line	differential regulation of protein folding and degradation as well as differential expression of other proteins likely to relate to CF pathology	(33)
human bronchial epithelial cell line	differential regulation of COP9 signalosome, cellular response to interleukin-4, response to salt stress, protein binding involved in protein folding, cell redox homeostasis & actin cytoskeleton	(34)
human nasal epithelial cells	impairment of pathways relating to metabolism, G protein process, inflammation and oxidative stress response, protein folding, proteolysis and structural proteins	(35)
human	differential enrichment/depletion of innate immunity,	(36)

bronchial epithelial cell line secretions	cytoskeleton/extracellular matrix organization & protease/antiprotease activity; oxidative & classical complement pathways altered *in vivo* were not altered *in vitro*	
cell-free sputum fluid or BALF	distinct CF protein profiles observed	(37-39)
cell-free BALF	less complete cohort clustering for young children with CF than for non-CF children with active lung disease	(40)
cell-free sputum fluid	2-D gel electrophoresis protein profiles from CF adults approached a “healthy” profile after hospitalized treatment	(38)
cell-free sputum fluid or BALF	increased immune response, proteolytic activity and dysregulation of complement in CF	(38, 39)
cell-free sputum fluid or BALF	identification of CF-associated inflammatory biomarkers and autoantigens, including myeloperoxidase, interleukin 8 and calgranulins A, B & C	(37, 38, 41)

Table E1: Summary of previous proteomic studies of the CF airways.

	IV abx for treatment of current PEx	long-term oral abx	oral abx (previous 30 days)	inhaled abx (previous 30 days)
CF1	meropenem & tobramycin	azithromycin	-	tobramycin & colistin
CF2	ceftazidime & tobramycin	-	-	tobramycin & colistin
CF3	cefuroxime & colistin	azithromycin	-	tobramycin
CF4	piperacillin/ tazobactam & tobramycin	azithromycin	-	colistin
CF5	piperacillin/ tazobactam, aztreonam & chlorphenamine	-	-	tobramycin & colistin
CF6	piperacillin/ tazobactam & tobramycin	azithromycin	ciprofloxacin	colistin
CF7	tobramycin & temocillin	azithromycin	-	tobramycin & colistin
CF11	piperacillin/ tazobactam & tobramycin	-	-	tobramycin
CF12	aztreonam & tobramycin	azithromycin	-	tobramycin & colistin
CF13	aztreonam & tobramycin	azithromycin	-	colistin
CF15	aztreonam & tobramycin	azithromycin	ciprofloxacin	colistin
CF16	tobramycin, ceftazidime & colistin	azithromycin	N	tobramycin

Table E2: Treatment regimens of the CF cohort.

UniProt Accession	Gene Name	Description	Log2 mean CF/control	Signif (*p*)
Q14764	MVP	Major vault protein	Reduced in CF	0.002
P28676	GCA	Grancalcin	2.347	0.000
P01833	PIGR	Polymeric immunoglobulin receptor	-1.467	0.000
O15144	ARPC2	Actin-related protein 2/3 complex subunit 2	-	n.s.
P52907	CAPZA1	F-actin-capping protein subunit alpha-1	-1.060	0.000
P05023	ATP1A1	Sodium/potassium-transporting ATPase subunit alpha-1	-	n.s.
P05091	ALDH2	Aldehyde dehydrogenase, mitochondrial	-1.944	0.000
Q9HD89	RETN	Resistin	1.968	0.000
P05141	SLC25A5	ADP/ATP translocase 2	-	n.s.
P38606	ATP6V1A	V-type proton ATPase catalytic subunit A	-1.586	0.000
P21281	ATP6V1B2	V-type proton ATPase subunit B, brain isoform	-1.153	0.000
Q01518	CAP1	Isoform 1 of Adenylyl cyclase-associated protein 1	-0.580	0.008
Q9UM07	PADI4	Protein-arginine deiminase type-4	-	n.s.
Q9Y6N5	SQRDL	Sulfide:quinone oxidoreductase, mitochondrial	-0.612	0.002
P13646	KRT13	Isoform 1 of Keratin, type I cytoskeletal 13	-3.704	0.000
P13667	PDIA4	Protein disulfide-isomerase A4	-0.798	0.000
P13796	LCP1	Plastin-2	0.362	0.049
P07237	P4HB	Protein disulfide-isomerase	0.302	0.027
P48735	IDH2	Isocitrate dehydrogenase [NADP], mitochondrial	0.551	0.012
P23368	ME2	NAD-dependent malic enzyme, mitochondrial	Increased in CF	0.000

P07339	CTSD	Cathepsin D	-2.120	0.000
P07384	CAPN1	Calpain-1 catalytic subunit	-0.864	0.001
P23528	CFL1	Cofilin-1	-	n.s.
P41218	MNDA	Myeloid cell nuclear differentiation antigen	1.051	0.000
P31943	HNRNPH1	Heterogeneous nuclear ribonucleoprotein H	-	n.s.
P15924	DSP	Isoform DPI of Desmoplakin	-2.154	0.000
Q15080	NCF4	Neutrophil cytosol factor 4	Increased in CF	0.000
Q15149	PLEC	Plectin	-1.452	0.000
P09622	DLD	Dihydrolipoyl dehydrogenase, mitochondrial	-0.930	0.000
Q6P4A8	PLBD1	Phospholipase B-like 1	-	n.s.
P51149	RAB7A	Ras-related protein Rab-7a	-0.379	0.022
P51159	RAB27A	Ras-related protein Rab-27A	1.698	0.000
P00367	GLUD1	Glutamate dehydrogenase 1, mitochondrial	-1.568	0.000
P00403	MT-CO2	Cytochrome c oxidase subunit 2	-0.759	0.006
P00450	CP	Ceruloplasmin	-0.662	0.016
P43353	ALDH3B1	Aldehyde dehydrogenase family 3 member B1	-0.878	0.001
P00505	GOT2	Aspartate aminotransferase, mitochondrial	-1.137	0.000
P43490	NAMPT	Nicotinamide phosphoribosyltransferase	0.797	0.000
P61626	LYZ	Lysozyme C	-0.548	0.002
P35579	MYH9	Myosin-9	-	n.s.
P27797	CALR	Calreticulin	-	n.s.
P27824	CANX	Calnexin	-	n.s.
P10606	COX5B	Cytochrome c oxidase subunit 5B, mitochondrial	-	n.s.
P10644	PRKAR1A	cAMP-dependent protein kinase type I-alpha regulatory subunit	0.914	0.000
Q92542	NCSTN	Nicastrin	-	n.s.

P20160	AZU1	Azurocidin	1.693	0.000
Q6UX06	OLFM4	Olfactomedin-4	1.636	0.000
P55072	VCP	Transitional endoplasmic reticulum ATPase	-1.079	0.000
P55084	HADHB	Trifunctional enzyme subunit beta, mitochondrial	-1.211	0.000
P12273	PIP	Prolactin-inducible protein	-1.746	0.000
Q08380	LGALS3BP	Galectin-3-binding protein	-2.038	0.000
O75131	CPNE3	Copine-3	1.068	0.000
P30040	ERP29	Endoplasmic reticulum resident protein 29	1.499	0.000
P12724	RNASE3	Eosinophil cationic protein	Increased in CF	0.001
P60842	EIF4A1	Eukaryotic initiation factor 4A-I	Increased in CF	0.004
P04792	HSPB1	Heat shock protein beta-1	-3.468	0.000
P04843	RPN1	Dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit 1 precursor	-1.065	0.000
P14314	PRKCSH	Glucosidase 2 subunit beta	-	n.s.
P14625	HSP90B1	Endoplasmin	-0.391	0.009
P40121	CAPG	Macrophage-capping protein	-	n.s.
O75955	FLOT1	Flotillin-1	0.782	0.001
P30740	SERPINB1	Leukocyte elastase inhibitor	1.050	0.000
P14780	MMP9	Matrix metalloproteinase-9	-	n.s.
P22894	MMP8	Neutrophil collagenase	1.777	0.000
P08311	CTSG	Cathepsin G	2.147	0.000
P61158	ACTR3	Actin-related protein 3	-0.648	0.000
O60234	GMFG	Glia maturation factor gamma	1.663	0.000
P50395	GDI2	Rab GDP dissociation inhibitor beta	-	n.s.
P40939	HADHA	Trifunctional enzyme subunit alpha, mitochondrial	-1.428	0.000

P00558	PGK1	Phosphoglycerate kinase 1	-	n.s.
P13639	EEF2	Elongation factor 2	-1.224	0.000
P10599	TXN	Thioredoxin	-1.289	0.000
P07737	PFN1	Profilin-1	-	n.s.
P30038	ALDH4A1	Delta-1-pyrroline-5-carboxylate dehydrogenase, mitochondrial	-	n.s.
P04839	CYBB	Cytochrome b-245 heavy chain	1.197	0.000
P09917	ALOX5	Arachidonate 5-lipoxygenase	-	n.s.
P04083	ANXA1	Annexin A1	-1.957	0.000
P52209	PGD	6-phosphogluconate dehydrogenase, decarboxylating	-	n.s.
P27105	STOM	Erythrocyte band 7 integral membrane protein	1.685	0.000
P09211	GSTP1	Glutathione S-transferase P	-0.741	0.000
P15144	ANPEP	Aminopeptidase N	0.486	0.007
Q8WUM4	PDCD6IP	Programmed cell death 6-interacting protein	-	n.s.
Q12913	PTPRJ	Receptor-type tyrosine-protein phosphatase eta precursor	0.621	0.006
P40926	MDH2	Malate dehydrogenase, mitochondrial	-1.000	0.000
Q8TDL5	BPIFB1	BPI fold-containing family B member 1	-1.470	0.003
Q14739	LBR	Lamin-B receptor	-	n.s.
P49913	CAMP	Cathelicidin antimicrobial peptide precursor	1.215	0.002
Q8IX19	MCEMP1	Mast cell-expressed membrane protein 1	Increased in CF	0.000
Q9Y2J8	PADI2	Protein-arginine deiminase type-2	0.915	0.005
P39656	DDOST	Dolichyl-diphosphooligosaccharide--protein glycosyltransferase 48 kDa subunit	-0.909	0.001
Q13576	IQGAP2	Ras GTPase-activating-like protein IQGAP2	0.546	0.001
P80188	LCN2	Neutrophil gelatinase-associated lipocalin	1.806	0.000

Q15084	PDIA6	Protein disulfide-isomerase A6	-0.570	0.035
P21333	FLNA	Isoform 2 of Filamin-A	-0.897	0.000
P50990	CCT8	T-complex protein 1 subunit theta	-0.833	0.000
P06576	ATP5B	ATP synthase subunit beta, mitochondrial	-1.416	0.000
A6NI72	NCF1B	Putative neutrophil cytosol factor 1B	Increased in CF	0.000
P10909	CLU	Clusterin	-2.856	0.000
P69905	HBA1 & HBA2	Hemoglobin subunit alpha	-	n.s.
P07355	ANXA2	Isoform 2 of Annexin A2	-2.744	0.000
P62937	PPIA	Peptidyl-prolyl cis-trans isomerase A	-	n.s.
P25705	ATP5A1	ATP synthase subunit alpha, mitochondrial	-0.751	0.010
P62805	HIST1H4F/ HIST1H4C; HIST1H4A; HIST1H4I; HIST1H4B; HIST2H4B; HIST1H4L; HIST1H4E; HIST1H4D; HIST1H4J; HIST1H4K; HIST1H4H; HIST2H4A; HIST4H4	Histone H4	1.775	0.000
Q14204	DYNC1H1	Cytoplasmic dynein 1 heavy chain 1	Reduced in CF	0.001
P06733	ENO1	Alpha-enolase	-0.694	0.002
P04040	CAT	Catalase	1.108	0.000

P14618	PKM	Pyruvate kinase PKM	-	n.s.
P13489	RNH1	Ribonuclease inhibitor	-0.549	0.001
P01011	SERPINA3	Alpha-1-antichymotrypsin	0.548	0.048
P01009	SERPINA1	Alpha-1-antitrypsin	1.036	0.001
P29401	TKT	Transketolase	0.955	0.000
P22314	UBA1	Ubiquitin-like modifier-activating enzyme 1	-	n.s.
P37837	TALDO1	Transaldolase	-0.581	0.019
O75083	WDR1	WD repeat-containing protein 1	-	n.s.
P01024	C3	Complement C3	Reduced in CF	0.020
Q14254	FLOT2	Flotillin-2	-	n.s.
P50995	ANXA11	Annexin A11	1.198	0.000
Q5JTV8	TOR1AIP1	Torsin-1A-interacting protein 1	-0.511	0.001
O43451	MGAM	Maltase-glucoamylase, intestinal	2.495	0.000
P36222	CHI3L1	Chitinase-3-like protein 1	CF only	n.a.
P61160	ACTR2	Actin-related protein 2	CF only	n.a.
Q9H9B4	SFXN1	Sideroflexin-1	CF only	n.a.
P48595	SERPINB10	Serpin B10	CF only	n.a.
Q13043	STK4	Serine/threonine-protein kinase 4	CF only	n.a.
P15498	VAV1	Proto-oncogene vav	CF only	n.a.
Q13636	RAB31	Ras-related protein Rab-31	CF only	n.a.
Q9H4M9	EHD1	EH domain-containing protein 1	CF only	n.a.
Q99829	CPNE1	Copine-1	CF only	n.a.
P20592	MX2	Interferon-induced GTP-binding protein Mx2	CF only	n.a.
P57737	CORO7	Coronin-7	CF only	n.a.
P34059	GALNS	N-acetylgalactosamine-6-sulfatase	CF only	n.a.
P26038	MSN	Moesin	CF only	n.a.

P08473	MME	Neprilysin	CF only	n.a.
Q9P107	GMIP	GEM-interacting protein	CF only	n.a.
P07332	FES	Tyrosine-protein kinase Fes/Fps	CF only	n.a.
Q9NUQ9	FAM49B	Protein FAM49B	CF only	n.a.
P98171	ARHGAP4	Rho GTPase-activating protein 4	CF only	n.a.
Q86YV0	RASAL3	RAS protein activator like-3	CF only	n.a.
Q14005	IL16	Pro-interleukin-16	CF only	n.a.
P30153	PPP2R1A	Serine/threonine-protein phosphatase 2A 65 kDa regulatory subunit A alpha isoform	CF only	n.a.
P62834	RAP1A	Ras-related protein Rap-1A	control only	n.a.
Q07954	LRP1	Prolow-density lipoprotein receptor-related protein 1	control only	n.a.
P02786	TFRC	Transferrin receptor protein 1	control only	n.a.
Q9NZM1	MYOF	Myoferlin	control only	n.a.
P10253	GAA	Lysosomal alpha-glucosidase	control only	n.a.
P07858	CTSB	Cathepsin B	control only	n.a.
Q5JXB2	UBE2NL	Putative ubiquitin-conjugating enzyme E2 N-like	control only	n.a.
Q96TC7	RMDN3	Regulator of microtubule dynamics protein 3	control only	n.a.
P13686	ACP5	Tartrate-resistant acid phosphatase type 5	control only	n.a.
Q5TD94	RSPH4A	Radial spoke head protein 4 homolog A	control only	n.a.
Q5RHP9	ERICH3	Glutamate-rich protein 3	control only	n.a.
P18124	RPL7	60S ribosomal protein L7	control only	n.a.
Q9NQ38	SPINK5	Serine protease inhibitor Kazal-type 5	control only	n.a.
Q8N392	ARHGAP1	Rho GTPase-activating protein 18	control only	n.a.

	8			
Q13838	DDX39B	Spliceosome RNA helicase DDX39B	control only	n.a.
P19013	KRT4	Keratin, type II cytoskeletal 4	control only	n.a.
O15031	PLXNB2	Plexin-B2	control only	n.a.

Table E3: Proteins detected in all 48 sputum samples or all samples belonging to at least one cohort. Where independent samples T-tests identified differential abundance between the cohorts (p<0.05) the $\log_2$ ratio of mean abundance levels and significance are recorded. Where Mann-Whitney U Tests identified differential abundance (p<0.05) increased or decreased abundance in CF is recorded along with significance. No significant difference in abundance between the cohorts is recorded as "-".

Functions Category	Bio Functions Annotation	Confidence level (*p*-value)	Regulation z-score
Cell Death	apoptosis of neutrophils	3.56E-09	-1.308
Cell Death	apoptosis of myeloid cells	6.85E-07	-1.108
Cell Death	killing of cells	6.31E-05	1.88
Cell Death	cell survival	3.15E-04	-1.014
Cell Death	cell death	3.42E-08	0.875
Cell Death	apoptosis of phagocytes	9.75E-07	-0.913
Cell Death	apoptosis	2.90E-06	0.648
Cell Death	cell death of lung cell lines	2.03E-05	
Cell Death	cell viability	1.41E-04	-0.895
Cell Death	cell death of muscle cells	2.36E-04	0.56
Cell Death	cell death of organ	2.63E-04	0.885
Cell Death	apoptosis of organ	2.86E-04	0.29
Cell Death	apoptosis of lung cell lines	3.34E-04	
Cell Death	apoptosis of muscle cells	3.45E-04	0.342
Cell-To-Cell Signaling & Interaction	adhesion of granulocytes	5.07E-04	1.465
Cell-To-Cell Signaling & Interaction	activation of cells	2.63E-03	1.06
Cell-To-Cell Signaling & Interaction	response of granulocytes	4.84E-04	-0.392
Cellular Assembly & Organization	disruption of lipid bilayer	1.88E-04	
Cellular Assembly & Organization	fusion of liposome	2.82E-04	

Cellular Compromise	injury of cells	1.06E-04	-0.261
Cellular Function & Maintenance	cellular homeostasis	5.25E-04	1.165
Cellular Function & Maintenance	autophagy	3.62E-04	0.607
Cellular Growth & Proliferation	proliferation of immune cells	3.68E-03	-2.182
Cellular Growth & Proliferation	proliferation of muscle cells	6.08E-03	1.581
Cellular Growth & Proliferation	proliferation of lymphocytes	8.17E-03	-1.975
Cellular Growth & Proliferation	growth of cells	4.81E-06	-0.016
Cellular Movement	migration of cells	2.71E-07	1.311
Cellular Movement	leukocyte migration	3.86E-06	1.82
Cellular Movement	cell movement of myeloid cells	1.10E-05	1.68
Cellular Movement	cell movement of granulocytes	2.36E-05	1.121
Cellular Movement	chemotaxis of cells	5.03E-05	1.142
Cellular Movement	invasion of cells	1.44E-04	1.001
Cellular Movement	cell movement of antigen presenting cells	2.31E-04	1.793
Cellular Movement	recruitment of leukocytes	3.66E-03	1.188
Cellular Movement	cell movement	5.53E-07	0.987
Cellular Movement	migration of granulocytes	3.17E-04	0.578
Free Radical Scavenging	metabolism of reactive oxygen species	1.85E-05	1.655
Free Radical Scavenging	generation of reactive oxygen species	4.71E-05	1.534
Free Radical Scavenging	synthesis of reactive oxygen	6.23E-05	1.729

	species		
Inflammatory Response	chemotaxis of leukocytes	1.03E-05	1.426
Inflammatory Response	cell movement of neutrophils	2.25E-05	2.036
Inflammatory Response	inflammatory response	2.92E-05	1.652
Inflammatory Response	cell movement of phagocytes	7.17E-05	2.272
Inflammatory Response	chemotaxis of granulocytes	8.05E-05	2.135
Inflammatory Response	chemotaxis of myeloid cells	8.29E-05	2.339
Inflammatory Response	chemotaxis of phagocytes	1.04E-04	1.641
Inflammatory Response	cell movement of macrophages	1.97E-04	1.877
Inflammatory Response	migration of neutrophils	9.41E-04	1.476
Inflammatory Response	chemotaxis of neutrophils	2.42E-03	2.284
Inflammatory Response	activation of myeloid cells	4.00E-03	1.195
Inflammatory Response	cell movement of monocytes	4.92E-03	1.481
Inflammatory Response	immune response	2.81E-06	0.689
Inflammatory Response	phagocytosis of granulocytes	4.41E-04	
Organismal Injury & Abnormalities	injury of organ	1.72E-05	-0.021
Post-Translational Modification	N-glycosylation of protein	7.48E-05	
Protein Synthesis	degradation of protein	1.02E-06	-0.653
Protein Synthesis	metabolism of protein	4.39E-06	-0.732
Protein Synthesis	damage of protein	1.88E-04	
Small Molecule Biochemistry	production of nitric oxide	7.49E-03	-1.821
Small Molecule Biochemistry	catabolism of acidic amino acid	3.93E-04	

Table E4: Ingenuity Pathway Analysis Bio Functions predicted by Fisher's Exact Test as over-represented in CF (p<0.05) with Regulation z-scores >1.0 or <-1.0 (i.e. indicating a greater than 2-fold difference). Shaded rows record Bio Functions with a high prediction confidence (p<0.0005), but where the predicted Regulation z-score indicates a less than 2-fold difference or none could be calculated.

UniProt Accession	Gene Name	Description	GO Biological Processes
P36222	CHI3L1	Chitinase-3-like protein 1	activation of NF-kappaB-inducing kinase activity; carbohydrate metabolic process; inflammatory response
P61160	ACTR2	Actin-related protein 2	cellular component movement
Q9H9B4	SFXN1	Sideroflexin-1	erythrocyte differentiation; cellular iron ion homeostasis
P48595	SERPINB10	Serpin B10	regulation of proteolysis
Q13043	STK4	Serine/threonine-protein kinase 4	induction of apoptosis; negative regulation of cell proliferation
P15498	VAV1	Proto-oncogene vav	induction of apoptosis by extracellular signals; immune response; regulation of small GTPase mediated signal transduction
Q13636	RAB31	Ras-related protein Rab-31	Golgi vesicle transport; protein transport; small GTPase mediated signal transduction
Q9H4M9	EHD1	EH domain-containing protein 1	GTP catabolic process; endocytic recycling; intracellular protein transport
Q99829	CPNE1	Copine-1	lipid metabolic process; vesicle-mediated transport
P20592	MX2	Interferon-induced GTP-binding protein Mx2	regulation of cell cycle; regulation of nucleocytoplasmic transport
P57737	CORO7	Coronin-7	Golgi membrane; cytoplasmic membrane-bounded vesicle
P34059	GALNS	N-acetylgalactosamine-6-sulfatase	carbohydrate metabolic process; keratan sulfate catabolic process
P26038	MSN	Moesin	leukocyte cell-cell adhesion; leukocyte migration; regulation of lymphocyte migration
P08473	MME	Neprilysin	beta-amyloid metabolic process; cellular response to cytokine stimulus; proteolysis
Q9P107	GMIP	GEM-interacting protein	negative regulation of Rho GTPase activity
P07332	FES	Tyrosine-protein kinase Fes/Fps	regulation of cell proliferation; positive regulation of myeloid cell differentiation; regulation of mast cell degranulation
Q9NUQ9	FAM49B	Protein FAM49B	-
P98171	ARHGAP4	Rho GTPase-activating protein 4	Rho protein signal transduction; induction of apoptosis by extracellular signals; cytoskeleton organization
Q86YV0	RASAL3	RAS protein activator like-3	negative regulation of Ras protein signal transduction
Q14005	IL16	Pro-interleukin-16	immune response; induction of positive chemotaxis; leukocyte chemotaxis
P30153	PPP2R1A	Serine/threonine-protein phosphatase 2A 65 kDa regulatory subunit A alpha isoform	inactivation of MAPK activity; induction of apoptosis; negative regulation of cell growth

P23368	ME2	NAD-dependent malic enzyme, mitochondrial	malate metabolic process
Q15080	NCF4	Neutrophil cytosol factor 4	immune response; phagosome maturation; antigen processing and presentation of exogenous peptide antigen via MHC class I, TAP-dependent; cell communication
P12724	RNASE3	Eosinophil cationic protein	RNA catabolic process; defense response to bacterium
Q8IX19	MCEMP1	Mast cell-expressed membrane protein 1	-
A6NI72	NCF1B	Putative neutrophil cytosol factor 1B	cell communication
P62834	RAP1A	Ras-related protein Rap-1A	activation of MAPKK activity; energy reserve metabolic process; small GTPase mediated signal transduction
Q13838	DDX39B	Spliceosome RNA helicase DDX39B	ATP-dependent helicase activity
Q5JXB2	UBE2NL	Putative ubiquitin-conjugating enzyme E2 N-like	-
Q5TD94	RSPH4A	Radial spoke head protein 4 homolog A	cilium movement
Q9NQ38	SPINK5	Serine protease inhibitor Kazal-type 5	epithelial cell differentiation; negative regulation of immune response; negative regulation of proteolysis
P10253	GAA	Lysosomal alpha-glucosidase	glucose metabolic process; glycogen catabolic process; lysosome organization
Q96TC7	RMDN3	Regulator of microtubule dynamics protein 3	apoptotic process; cell differentiation
P19013	KRT4	Keratin, type II cytoskeletal 4	cytoskeleton organization; epithelial cell differentiation; negative regulation of epithelial cell proliferation
P07858	CTSB	Cathepsin B	autophagy; negative regulation of cell death; regulation of apoptotic process; proteolysis
Q5RHP9	ERICH3	Glutamate-rich protein 3	-
Q8N392	ARHGAP18	Rho GTPase-activating protein 18	positive regulation of GTPase activity
Q9NZM1	MYOF	Myoferlin	plasma membrane repair; regulation of vascular endothelial growth factor receptor signaling pathway
P02786	TFRC	Transferrin receptor protein 1	cellular iron ion homeostasis; proteolysis
P18124	RPL7	60S ribosomal protein L7	structural constituent of ribosome; transcription regulator activity; translation
O15031	PLXNB2	Plexin-B2	positive regulation of GTPase activity
P13686	ACP5	Tartrate-resistant acid phosphatase type 5	response to LPS; negative regulation of inflammatory response; negative regulation of nitric oxide biosynthetic process;

			negative regulation of superoxide anion generation
Q07954	LRP1	Prolow-density lipoprotein receptor-related protein 1	apoptotic cell clearance; beta-amyloid clearance; regulation of phospholipase A2 activity; regulation of actin cytoskeleton organization
Q14764	MVP	Major vault protein	cell proliferation; negative regulation of protein autophosphorylation; negative regulation of signaling; protein activation cascade; protein transport
Q14204	DYNC1H1	Cytoplasmic dynein 1 heavy chain 1	antigen processing and presentation of exogenous peptide antigen via MHC class II; cell death; microtubule-based movement; stress granule assembly
P01024	C3	Complement C3	complement activation, alternative & classical pathways; inflammatory response; negative regulation of endopeptidase activity; positive regulation of activation of membrane attack complex; positive regulation of angiogenesis; positive regulation of glucose transport; positive regulation of lipid storage; positive regulation of phagocytosis; regulation of triglyceride biosynthetic process

Table E5: Gene Ontology Biological Processes for proteins detected exclusively in the CF cohort (dark red) or the control cohort (dark blue) and proteins up-regulated (pale red) or down-regulated (pale blue) in CF relative to controls (Mann-Whitney U test; $p<0.05$).

Run	Samples
1	CF1.1, CF1.2, CF1.3, CF2.1, CF2.2, CF2.3 & REF
2	CF3.1, CF3.2, CF3.3, CF6.1, CF6.2, CF6.3 & REF
3	CF4.1, CF4.2, CF4.3, CF5.1, CF5.2, CF5.3 & REF
4	CF7.1, CF7.2, CF7.3, CF11.1, CF11.2, CF11.3 & REF
5	CF12.1, CF12.2, CF12.3, CF13.1, CF13.2, CF13.3 & REF
6	CF15.1, CF15.2, CF15.3, CF16.1, CF16.2, CF16.3 & REF
7	HC23, HC24, HC33, HC35, HC36, HC38 & REF
8	HC27, HC28, HC29, HC39, HC40, HC41 & REF

Table E6: Samples included per MudPIT run. CF indicates the CF cohort and HC the healthy control cohort. The code used to label the CF samples indicates first patient ID followed by ".1" for exacerbation pre-antibiotic treatment, ".2" for exacerbation post-antibiotic treatment (<24h) and ".3" for stable follow-up (4-15 weeks following treatment.) Only one sample was collected from each of the health controls.

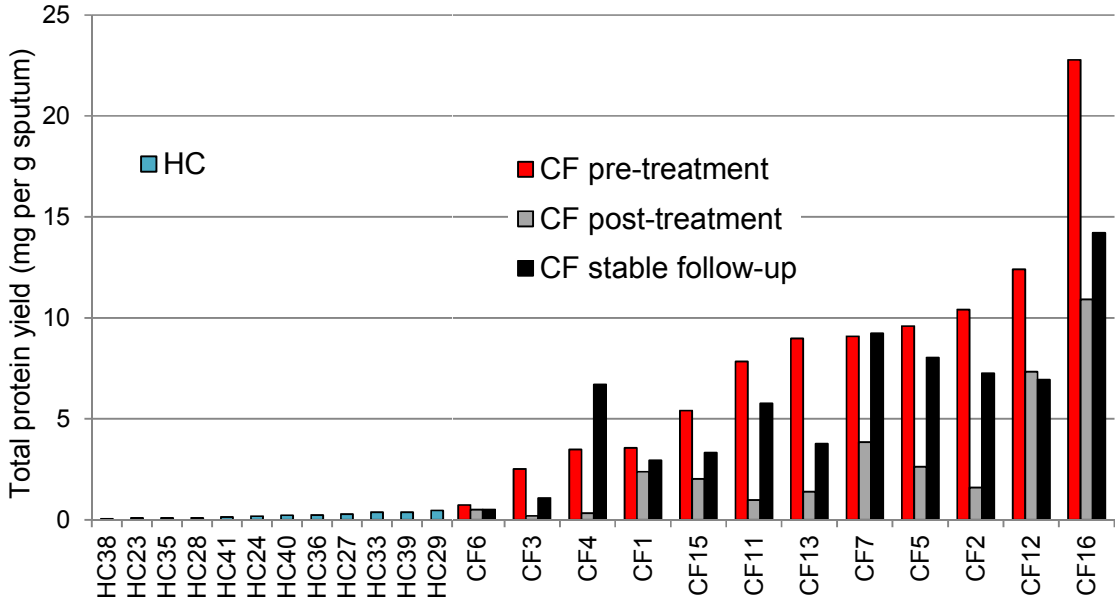


Figure E1: Sputum total protein yields (mg per g sputum) for single samples collected from healthy controls (HC) and for three samples collected from each individual in the CF cohort. The three CF samples were collected: (1) at exacerbation before commencement of antibiotic treatment; (2) at completion of antibiotic treatment for exacerbation; and (3) during a follow-up period of clinical stability.

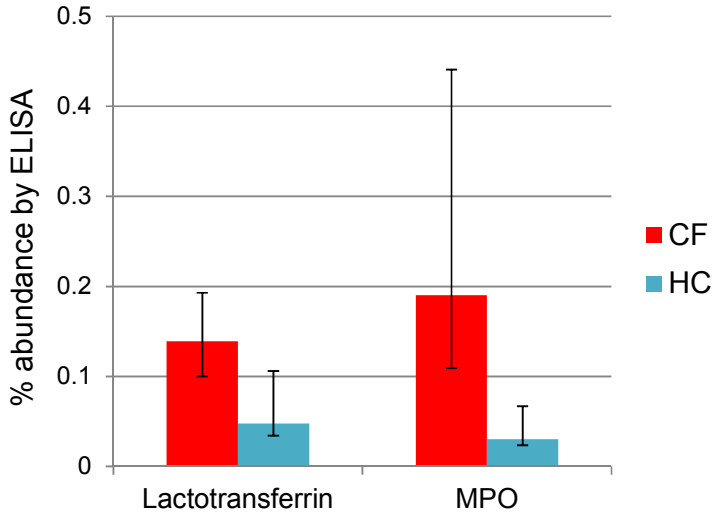


Figure E2: Median (±IQR) % abundance of myeloperoxidase and lactotransferrin per sample as determined by ELISA quantification. For both proteins, median % abundance is greater in the CF cohort compared to the healthy control (HC) cohort (p<0.05, Mann Whitney U test).

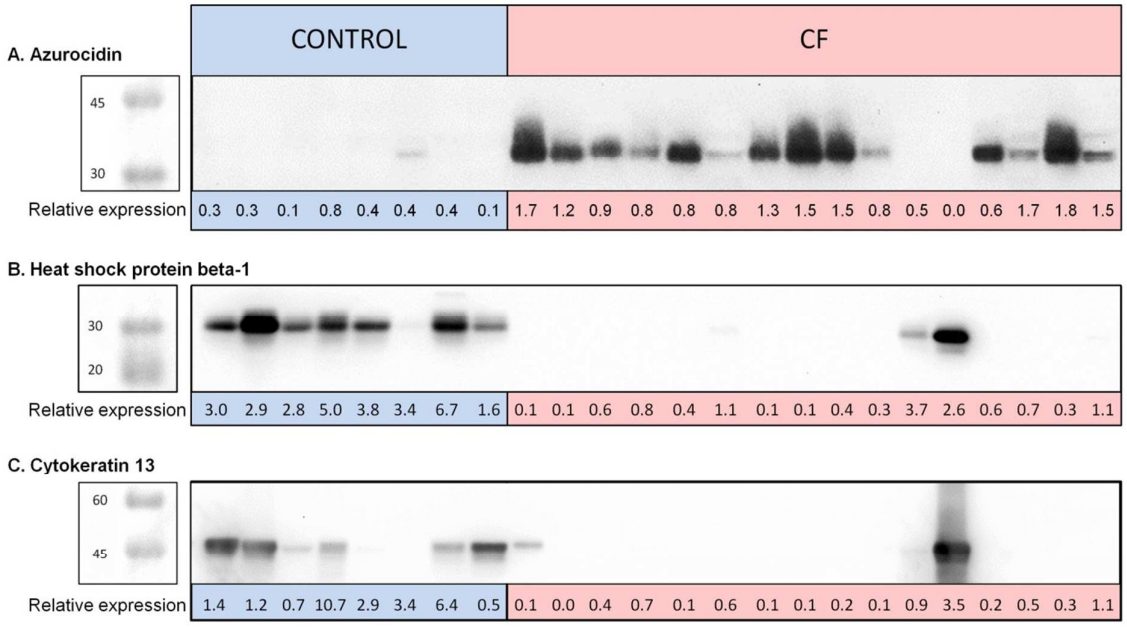


Figure E3: Western Blot validation of mass spectrometry data for azurocidin, heat shock protein beta-1 and cytokeratin 13. Relative abundance values were calculated as % emPAI from the mass spectrometry data.

Proteomic profile of cystic fibrosis sputum cells in adults chronically infected with *Pseudomonas aeruginosa*

Sally H. Pattison, David S. Gibson, Elinor Johnston, Samantha Peacock, Keith Rivera, Michael M. Tunney, Darryl J. Pappin, and J. Stuart Elborn

Supplementary Data

METHODS

Extraction of sputum cellular proteins

All work was carried out at 4℃ and using prechilled buffers. Approximately 1g of mucus plug was separated from each sputum sample and mixed with 9 ml sterile phosphate buffered saline pH 7.4 (PBS) containing Roche Complete Mini protease inhibitor cocktail (PIC). The mixture was homogenised thoroughly by repeat syringing (no needle) and cells pelleted by centrifugation for 10 min at 4,000 x g. The cells were washed twice more in 10 ml PBS+PIC and strained through a 70 µm nylon cell strainer (BD Biosciences) to remove any remaining cell clumps. The cells were pelleted once more to remove all supernatant before protein extraction. Cell pellets were resuspended in 0.2 - 3.0 ml lysis buffer (10 mM HEPES, 0.5 mM EDTA, 1% v/v Nonidet P 40, 4% w/v SDS, pH 8.0) depending on pellet size and Sigma protease inhibitor cocktail added at 10 µl per 1 ml as directed. After thorough homogenisation by vortexing and repeat passage through an 18-gauge needle, the mixtures were subjected to 5 freeze-thaw cycles and water bath sonication to ensure lysis. Insoluble cell debris was removed by centrifugation at 14,000 x g and the soluble protein supernatants were stored at -80℃.

Protein digestion and iTRAQ labelling

Protein digestion and iTRAQ labelling were performed as described previously (1). Protein concentrations were determined by BCA Protein Assay (Pierce). ProteaseMax (Promega) made up in 50 mM triethylammonium bicarbonate (TEAB) was added to a final concentration of 0.1% to 70 µg of each protein sample. The samples were reduced with 5 mM tris-(2-carboxyethyl) phosphine at 55℃ for 20 min and alkylated with 10 mM methyl methanethiosulphonate at room temperature for 20 min. Following methanol/chloroform precipitation for 60 min at 4℃, protein pellets

were resuspended in 50 mM TEAB, 0.1% Proteasemax and trypsin digested overnight at 37°C. The tryptic peptides were dried to <5 µl by vacuum centrifugation and diluted with 500 mM TEAB to a final volume of 30 µl, before acidification with trifluoroacetic acid and labelling with 8plex iTRAQ Reagents (AB Sciex) according to kit instructions.

MudPIT

Six test samples and a reference sample were mixed for simultaneous analysis by 2D liquid chromatography MudPIT, using a two-dimensional Vented Column Setup with a Proxeon nano-flow high-performance liquid pump as described previously (1). Eight MudPIT runs were performed in total and the samples combined per run are recorded in Table E6. Each salt step was eluted at 3.5 µl/min on an analytical column (100 µm internal diameter packed with 12 cm of 5 µm Aqua C18) on line with an LTQ Orbitrap Velos (Thermo Electron). Data were acquired in profile mode using the following parameters: for full-scan Fourier transform mass spectrometry, resolution = 60,000, m/z = 380–1,700 and the 10 most intense ions were fragmented with higher-collision dissociation at a normalised collision energy of 40% and an activation time of 0.1. Minimum threshold signal was at 5,000 and isolation width at 1.2. Dynamic exclusion settings were repeat count 1, repeat duration of 30, exclusion list size 500, exclusion duration 60 and exclusion mass width 10 p.p.m.

Protein identification

Peaklists were generated by Mascot Distiller (Matrix science; version 2.3). Protein identification was performed by the Mascot search engine version 2.3 against the UniProt human database 2016 (92,179 sequences) plus NCBI non-redundant database (2016), taxonomy *P. aeruginosa* (634,627 sequences) with a precursor mass tolerance of 20 p.p.m., with iTRAQ8plex (N-terminal) and iTRAQ8plex (lysine)

as fixed modifications and oxidation (methionine) and deamidated (asparagine & lysine) as variable modifications. Trypsin was set as the cleavage enzyme and one missed cleavage was allowed. A mass tolerance of 0.3 Da was set for fragment ions. Mascot Percolator was used to re-score the results to false discovery rates of 0.74-1.16% for protein identification. Protein identifications with a MOWSE score less than 65 were rejected. Protein-level iTRAQ ratios were calculated as intensity weighted, using only peptides with expectation values less than 0.05. Global ratio normalization was performed using intensity summation, with no outlier rejection.

Enzyme-linked Immunosorbent assay (ELISA)

In order to validate myeloperoxidase and lactotransferrin as high abundance proteins in the CF cohort compared to control cohort and to demonstrate that commercial ELISAs are a practical method for validation of candidate biomarkers, both proteins were quantified in each sample by ELISA (ab119605 and ab108882 respectively, Abcam, Cambridge, UK) according to the manufacturer's instructions. Figure E2 shows the median±IQR % abundance for each candidate biomarker along with the corresponding data for the control cohort.

Western immunoblot

Western immunoblot was used to further confirm the mass spectrometry identification and quantification data (Figure E3). Patient sample protein concentrations were assessed by BCA assay (Thermo Scientific Inc., Rockford, USA). 25 µg of protein was loaded per lane on 4-20% TGX™ gels and transferred to 0.45 µm PVDF membrane (Millipore Corproration, Massachusetts, USA) with a Criterion™ mini gel tank and blot module, associated buffers and power supply, according to the manufacturer's guidelines (Bio-rad Laboratories Inc., Hercules, USA). Nonspecific binding was blocked overnight at 4 ℃ using 5% non-fat milk in PBS with 0.05% Tween 20 (PBST) (Sigma-Aldrich Company Ltd., Gillingham, UK). Primary antibodies against azurocidin, heatshock protein beta 1 and cytokeratin (Abcam Ltd, Cambridge, UK) were incubated with blotted membranes at a diluted in PBST, for 1 h at room temperature. Membranes were washed 3 times (5 min each) with PBST and incubated in secondary anti-IgG horseradish peroxidase diluted in PBST (Promega Inc., Madison, USA), for 1 h at room temperature. After a further 6 washes, bands were visualized using Supersignal West Pico chemiluminescent substrate (Thermo Scientific Inc., Rockford, USA), imaged with a Autochemi CCD camera and analysed with Labworks software version 4.0.0.8 (UVP Ltd., Cambridge, UK).

DISCUSSION

A number of proteins detected consistently in CF sputum correlated negatively with lung function measured as FEV_1 % predicted. Their biological/clinical relevance and potential for causing injury are discussed below. Interestingly, 9 out of these 15 proteins also showed differential abundance in the CF cohort following intravenous antibiotic treatment of a pulmonary exacerbation, with the changes in abundance bringing them closer to levels in the control cohort.

Acute inflammation-associated proteins alpha-1-antitrypsin, haptoglobin, fibrinogen and alpha-2-macroglobulin have previously been identified as serum biomarkers of pulmonary disease including CF and COPD and can be predictive of response to treatment and prognosis (2-6). The presence of these proteins in the lungs has traditionally been thought to be due to plasma leakage. However, haptoglobin and fibrinogen may also be synthesized in the lungs in response to inflammation and injury, and may exert roles in immunity and repair mechanisms (7, 8). Alpha-1-antitrypsin is an inhibitor of neutrophil serine proteases, including elastase, which contribute significantly to lung destruction in CF, and therefore its negative correlation with lung function might seem surprising. However, previous studies observed that plasma deficiencies during inflammation were associated with less severe lung disease in CF (9, 10). Fibrinogen has been shown to increase mucin production in airway epithelial cells (11), which could contribute to the CF hypersecretory phenotype.

Three proteins associated with the migration of neutrophils, lymphocytes and alveolar monocytes – integrin beta 2, vanin-2 and serine/threonine-protein kinase 10 (12-14) – correlated negatively with lung function. This correlation could be due to either the negative impact of these cells as part of an overactive immune system or

poor lung health stimulating an increased influx of immune cells. In accordance with the predicted neutrophil influx in CF, four granule proteins – neutrophil elastase, olfactomedin-4, CD63 antigen and maltase-glucoamylase – were also found to correlate negatively with lung function. Neutrophil elastase exerts antimicrobial activity and therefore might be expected to promote lung health. However, CF-associated accumulation of neutrophil elastase can be damaging to lung tissue (15-17) and sputum measurements have previously been shown to predict the rate of FEV_1 decline in children with CF (18), while higher levels of neutrophil elastase following antibiotic treatment of an exacerbation have been associated with an increased risk of subsequent exacerbation (19). Additionally, serum levels of elastase/alpha-1-antitrypsin complex have been found to predict response to treatment of chronic *P. aeruginosa* colonisation (3). In contrast, olfactomedin-4 has been shown to down-regulate proinflammatory responses to bacterial infection and reduce bacterial killing mediated by granule serine proteinases (20, 21), which would be expected to negatively affect lung health. Additionally, olfactomedin-4 has recently been identified as a novel target autoantigen for antineutrophil cytoplasmic antibodies (ANCA), although both cases reported were for people without CF (22).

Neprilysin cleaves inflammatory peptides and has been proposed as protective against hypoxia- and smoke-related pulmonary vascular remodeling in COPD (23, 24), while thioredoxin-dependent peroxidase reductase has been implicated as an important scavenger of reactive oxygen species during LPS-induced oxidative stress and gene knockout mice showed increased lung inflammation (25). Although both correlated negatively with lung function, there is no published data suggesting their effects can be harmful. Thus, it is likely that observed levels signify a response to inflammation rather than a cause. BASP1 is strongly expressed in ciliated respiratory

cells (26) and its detection in sputum cells is likely indicative of significant dislodgement of these cells and injury to the respiratory tract that would be expect to hinder mucus clearance. Additionally, increased BASP1 expression has been linked to apoptosis of human tubular epithelial cells and in diabetic nephropathy (27). Finally, levels of RhoA also correlated negatively with lung function. As already discussed, RhoA is involved in the regulation of many activities relating to inflammatory and immune response repair such as phagocytosis, cell migration and wound closure. In part its negative correlation with lung function may reflect the degree of injury, but also its promotion of neutrophil influx may incur a mechanism for tissue damage.

REFERENCES

1 Scuoppo C, Miething C, Lindqvist L, Reyes J, Ruse C, Appelmann I, Yoon S, Krasnitz A, Teruya-Feldstein J, Pappin D, *et al.* A tumour suppressor network relying on the polyamine-hypusine axis. Nature 2012;487:244-248.

2 Engström G, Segelstorm N, Ekberg-Aronsson M, Nilsson PM, Lindgärde F, Löfdahl C. Plasma markers of inflammation and incidence of hospitalisations for COPD: results from a population-based cohort study. Thorax 2009;64:211-215.

3 Valletta EA, Rigo A, Bonazzi L, Zanolla L, Mastella G. Modification of some markers of inflammation during treatment for acute respiratory exacerbation in cystic fibrosis. Acta Paediatr 1992;81:227-230.

4 Duvoix A, Dickens J, Haq I, Mannino D, Miller B, Tal-Singer R, Lomas DA. Blood fibrinogen as a biomarker of chronic obstructive pulmonary disease. Thorax 2013;68:670-676.

5 Jiang R, Burke GL, Enright PL, Newman AB, Margolis HG, Cushman M, Tracy RP, Wang Y, Kronmal RA, Barr RG. Inflammatory markers and longitudinal lung function decline in the elderly. Am J Epidemiol 2008;168:602-610.

6 Sagel SD, Thompson V, Chmiel JF, Montgomery GS, Nasr.S Z, Perkett E, Saavedra MT, Slovis B, Anthony MM, Emmett P, *et al.* Effect of treatment of cystic fibrosis pulmonary exacerbations on systemic inflammation. Ann Am Thorac Soc 2015;12:708-717.

7 Abdullah M, Kähler D, Vock C, Reiling N, Kugler C, Drömann D, Rupp J, Hauber HP, Fehrenbach H, Zabel P, *et al.* Pulmonary haptoglobin and CD163 are functional immunoregulatory elements in the human lung. Respiration 2012;83:61-73.

8 Guadiz G, Sporn LA, Goss RA, Lawrence SO, Marder VJ, Simpson-Haidaris PJ. Polarized secretion of fibrinogen by lung epithelial cells. Am J Respir Cell Mol Biol 1997;17:60-69.

9 Mahadeva R, Westerbeek R, Perry D, Lovegrove J, Whitehouse D, Carroll N, Ross-Russell R, Webb A, Bilton D, Lomas D. Alpha1-antitrypsin deficiency alleles and the Taq-I G-->A allele in cystic fibrosis lung disease. Eur Respir J 1998;11:873-879.

10 Mahadeva R, Stewart S, Bilton D, Lomas DA. Alpha-1 antitrypsin deficiency alleles and severe cystic fibrosis lung disease. Thorax 1998;53:1022-1024.

11 Kim S, Nadel JA. Fibrinogen binding to ICAM-1 promotes EGFR-dependent mucin production in human airway epithelial cells. Am J Physiol Lung Cell Mol Physiol 2009;297:L174-183.

12 Rosseau S, Selhorst J, Wiechmann K, Leissner K, Maus U, Mayer K, Grimminger F, Seeger W, Lohmeyer J. Monocyte migration through the alveolar epithelial barrier: adhesion molecule mechanisms and impact of chemokines. J Immunol 2000;164:427-435.

13 Suzuki K, Watanabe T, Sakurai S, Ohtake K, Kinoshita T, Araki A, Fujita T, Takei H, Takeda Y, Sato Y, *et al.* A novel glycosylphosphatidyl inositol-anchored protein on human leukocytes: a possible role for regulation of neutrophil adherence and migration. J Immunol 1999;162:4277-4284.

14 Belkina NV, Liu Y, Hao J, Karasuyama H, Shaw S. LOK is a major ERM kinase in resting lymphocytes and regulates cytoskeletal rearrangement through ERM phosphorylation. Proc Natl Acad Sci USA 2009;106:4707-4712.

15 Amitani R, Wilson R, Rutman A, Read R, Ward C, Burnett D, Stockley RA, Cole PJ. Effects of human neutrophil elastase and Pseudomonas aeruginosa proteinases on human respiratory epithelium. Am J Respir Cell Mol Biol 1991;4:26-32.

16 O'Connor CM, Gaffney K, Keane J, Southey A, Byrne N, O'Mahoney S, Fitzgerald MX. alpha 1-Proteinase inhibitor, elastase activity, and lung disease severity in cystic fibrosis. Am Rev Respir Dis 1993;148:1665-1670.

17 Gehrig S, Duerr J, Weitnauer M, Wagner CJ, Graeber SY, Schatterny J, Hirtz S, Belaaouaj A, Dalpke AH, Schultz C, *et al.* Lack of neutrophil elastase reduces inflammation, mucus hypersecretion, and emphysema, but not mucus obstruction, in mice with cystic fibrosis-like lung disease. Am J Respir Crit Care Med 2014;189:1082-1092.

18 Sagel SD, Wagner BD, Anthony MM, Emmett P, Zemanick ET. Sputum biomarkers of inflammation and lung function decline in children with cystic fibrosis. Am J Respir Crit Care Med 2012;186:857-865.

19 Waters VJ, Stanojevic S, Sonneveld N, Klingel M, Grasemann H, Yau YCW, Tullis E, Wilcox P, Freitag A, Chilvers M, *et al.* Factors associated with response to treatment of pulmonary exacerbations in cystic fibrosis patients [published online ahead of print February 14, 2015]. J Cyst Fibros ;doi: 10.1016/j.jcf.2015.01.007.

20 Liu W, Yan M, Liu Y, Wang R, Li C, Deng C, Singh A, Coleman WG, Rodgers GP. Olfactomedin 4 down-regulates innate immunity against Helicobacter pylori infection. Proc Natl Acad Sci USA 2010;107:11056-11061.

21 Liu W, Yan M, Liu Y, McLeish KR, Coleman WG, Rodgers GP. Olfactomedin 4 inhibits cathepsin C-mediated protease activities, thereby modulating neutrophil killing of Staphylococcus aureus and Escherichia coli in mice. J Immunol 2012;189:2460-2467.

22 Amirbeagi F, Thulin P, Pullerits R, Pedersen B, Andersson BA, Dahlgren C, Welin A, Bylund J. Olfactomedin-4 autoantibodies give unusual c-ANCA staining patterns with reactivity to a subpopulation of neutrophils. J Leukoc Biol 2015;97:181-189.

23 Wick MJ, Buesing EJ, Wehling CA, Loomis ZL, Cool CD, Zamora MR, Miller YE, Colgan SP, Hersh LB, Voelkel NF, *et al.* Decreased neprilysin and pulmonary vascular remodeling in chronic obstructive pulmonary disease. Am J Respir Crit Care Med 2011;183:330-340.

24 Hashimoto S, Amaya F, Oh-Hashi K, Kiuchi K, Hashimoto S. Expression of neutral endopeptidase activity during clinical and experimental acute lung injury. Respir Res 2010;11:164.

25 Li L, Shoji W, Takano H, Nishimura N, Aoki Y, Takahashi R, Goto S, Kaifu T, Takai T, Obinata M. Increased susceptibility of MER5 (peroxiredoxin III) knockout mice to LPS-induced oxidative stress. Biochem Biophys Res Commun 2007;355:715-721.

26 Uhlén M, Björling E, Agaton C, Szigyarto CA, Amini B, Andersen E, Andersson A, Angelidou P, Asplund A, Asplund C, *et al.* A human protein atlas for normal and cancer tissues based on antibody proteomics. Mol Cell Proteomics 2005;4:1920-1932.

27 Sanchez-Niño MD, Sanz AB, Lorz C, Gnirke A, Rastaldi MP, Nair V, Egido J, Ruiz-Ortega M, Kretzler M, Ortiz A. BASP1 promotes apoptosis in diabetic nephropathy. J Am Soc Nephrol 2010;21:610-621.

28 Bensalem N, Ventura AP, Vallée B, Lipecka J, Tondelier D, Davezac N, Santos AD, Perretti M, Fajac A, Sermet-Gaudelus I, *et al.* Down-regulation of the anti-inflammatory protein annexin A1 in cystic fibrosis knock-out mice and patients. Mol Cell Proteomics 2005;4:1591-1601.

29 Roxo-Rosa M, da Costa G, Luider T, Scholte B, Coelho A, Amaral M, Penque D. Proteomic analysis of nasal cells from cystic fibrosis patients and non-cystic fibrosis control individuals: search for novel biomarkers of cystic fibrosis lung disease. Proteomics 2006;6:2314-2325.

30 Carvalho-Oliveira I, Charro N, Aarbiou J, Buijs-Offerman R, Wilke M, Schettgen T, Kraus T, Titulaer M, Burgers P, Luider T, *et al.* Proteomic analysis of naphthalene-induced airway epithelial injury and repair in a cystic fibrosis mouse model. J Proteome Res 2009;8:3606-3616.

31 Gomes-Alves P, Imrie M, Gray RD, Nogueira P, Ciordia S, Pacheco P, Azevedo P, Lopes C, de Almeida AB, Guardiano M, *et al.* SELDI-TOF biomarker signatures for cystic fibrosis, asthma and chronic obstructive pulmonary disease. Clin Biochem 2010;43:168-177.

32 Ciavardelli D, D'Orazio M, Pieroni L, Consalvo A, Rossi C, Sacchetta P, Di Ilio C, Battistoni A, Urbani A. Proteomic and ionomic profiling reveals significant alterations of protein expression and calcium homeostasis in cystic fibrosis cells. Mol Biosyst 2013;9:1117-1126.

33 Ruaniyar N, Gupta V, Balch WE, Yates JR3. Quantitative proteomic profiling reveals differentially regulated protein in cystic fibrosis cells. J Proteome Res 2014;13:4668-4675.

34 Lavallée-Adam M, Rauniyar N, McClatchy DB, Yates JR3. PSEA-Quant: a protein set enrichment analysis on label-free and label-based protein quantification data. J Proteome Res 2014;13:5496-5509.

35 Jeanson L, Guerrera IC, Papon JF, Chhuon C, Zadigue P, Prulière-Escabasse V, Amselem S, Escudier E, Coste A, Edelman A. Proteomic analysis of nasal epithelial cells from cystic fibrosis patients. PLoS One 2014;9:e108671.

36 Peters-Hall JR, Brown KJ, Pillai DK, Tomney A, Garvin LM, Wu X, Rose MC. Quantitative proteomics reveals an altered cystic fibrosis in vitro bronchial epithelial secretome. Am J Respir Cell Mol Biol 2015;53:22-32.

37 Gray RD, MacGregor G, Noble D, Imrie M, Dewar M, Boyd AC, Innes JA, Porteous DJ, Greening AP. Sputum proteomics in inflammatory and suppurative respiratory diseases. Am J Respir Crit Care Med 2008;178:444-452.

38 Sloane AJ, Lindner RA, Prasad SS, Sebastian LT, Pedersen SK, Robinson M, Bye PT, Nielson DW, Harry JL. Proteomic analysis of sputum from adults and children with cystic fibrosis and from control subjects. Am J Respir Crit Care Med 2005;172:1416-1426.

39 Gharib S, Vaisar T, Aitken M, Park D, Heinecke J, Fu X. Mapping the lung proteome in cystic fibrosis. J Proteome Res 2009;8:3020-3029.

40 MacGregor G, Gray RD, Hilliard TN, Imrie M, Boyd AC, Alton EWFW, Bush A, Davies JC, Innes JA, Porteous DJ, *et al.* Biomarkers for cystic fibrosis lung disease: Application of SELDI-TOF mass spectrometry to BAL fluid. J Cyst Fibros 2008;7:352-358.

41 Pedersen SK, Sloane AJ, Prasad SS, Sebastian LT, Lindner RA, Hsu M, Robinson M, Bye PT, Weinberger RP, Harry JL. An immunoproteomic approach for identification of clinical biomarkers for monitoring disease. Mol Cell Proteomics 2005;4:1052-1060.

Sample type	CF phenotype/observation	Refs
human nasal epithelial cells	reduced expression of the anti-inflammatory protein annexin A1 in CF	(28)
human nasal epithelial cells	differential expression of proteins related to chronic inflammation, oxidative stress, cytoskeleton proteins and mitochondrial proteins	(29)
mouse airway epithelial cells	reduced expression of carbonyl reductase (prostaglandin metabolism) and a more pronounced down-regulation of two key enzymes of retinoic acid metabolism after injury	(30)
human nasal epithelial cells	multiple-biomarker signature patterns for CF	(31)
human bronchial epithelial cell line	differences in glycolytic and gluconeogenic pathways, unfolded protein responses and a signaling pathway involved in cellular calcium homeostasis	(32)
human bronchial epithelial cell line	differential regulation of protein folding and degradation as well as differential expression of other proteins likely to relate to CF pathology	(33)
human bronchial epithelial cell line	differential regulation of COP9 signalosome, cellular response to interleukin-4, response to salt stress, protein binding involved in protein folding, cell redox homeostasis & actin cytoskeleton	(34)
human nasal epithelial cells	impairment of pathways relating to metabolism, G protein process, inflammation and oxidative stress response, protein folding, proteolysis and structural proteins	(35)
human	differential enrichment/depletion of innate immunity,	(36)

bronchial epithelial cell line secretions	cytoskeleton/extracellular matrix organization & protease/antiprotease activity; oxidative & classical complement pathways altered *in vivo* were not altered *in vitro*	
cell-free sputum fluid or BALF	distinct CF protein profiles observed	(37-39)
cell-free BALF	less complete cohort clustering for young children with CF than for non-CF children with active lung disease	(40)
cell-free sputum fluid	2-D gel electrophoresis protein profiles from CF adults approached a “healthy” profile after hospitalized treatment	(38)
cell-free sputum fluid or BALF	increased immune response, proteolytic activity and dysregulation of complement in CF	(38, 39)
cell-free sputum fluid or BALF	identification of CF-associated inflammatory biomarkers and autoantigens, including myeloperoxidase, interleukin 8 and calgranulins A, B & C	(37, 38, 41)

Table E1: Summary of previous proteomic studies of the CF airways.

	IV abx for treatment of current PEx	long-term oral abx	oral abx (previous 30 days)	inhaled abx (previous 30 days)
CF1	meropenem & tobramycin	azithromycin	-	tobramycin & colistin
CF2	ceftazidime & tobramycin	-	-	tobramycin & colistin
CF3	cefuroxime & colistin	azithromycin	-	tobramycin
CF4	piperacillin/ tazobactam & tobramycin	azithromycin	-	colistin
CF5	piperacillin/ tazobactam, aztreonam & chlorphenamine	-	-	tobramycin & colistin
CF6	piperacillin/ tazobactam & tobramycin	azithromycin	ciprofloxacin	colistin
CF7	tobramycin & temocillin	azithromycin	-	tobramycin & colistin
CF11	piperacillin/ tazobactam & tobramycin	-	-	tobramycin
CF12	aztreonam & tobramycin	azithromycin	-	tobramycin & colistin
CF13	aztreonam & tobramycin	azithromycin	-	colistin
CF15	aztreonam & tobramycin	azithromycin	ciprofloxacin	colistin
CF16	tobramycin, ceftazidime & colistin	azithromycin	N	tobramycin

Table E2: Treatment regimens of the CF cohort.

UniProt Accession	Gene Name	Description	Log2 mean CF/control	Signif (*p*)
Q14764	MVP	Major vault protein	Reduced in CF	0.002
P28676	GCA	Grancalcin	2.347	0.000
P01833	PIGR	Polymeric immunoglobulin receptor	-1.467	0.000
O15144	ARPC2	Actin-related protein 2/3 complex subunit 2	-	n.s.
P52907	CAPZA1	F-actin-capping protein subunit alpha-1	-1.060	0.000
P05023	ATP1A1	Sodium/potassium-transporting ATPase subunit alpha-1	-	n.s.
P05091	ALDH2	Aldehyde dehydrogenase, mitochondrial	-1.944	0.000
Q9HD89	RETN	Resistin	1.968	0.000
P05141	SLC25A5	ADP/ATP translocase 2	-	n.s.
P38606	ATP6V1A	V-type proton ATPase catalytic subunit A	-1.586	0.000
P21281	ATP6V1B2	V-type proton ATPase subunit B, brain isoform	-1.153	0.000
Q01518	CAP1	Isoform 1 of Adenylyl cyclase-associated protein 1	-0.580	0.008
Q9UM07	PADI4	Protein-arginine deiminase type-4	-	n.s.
Q9Y6N5	SQRDL	Sulfide:quinone oxidoreductase, mitochondrial	-0.612	0.002
P13646	KRT13	Isoform 1 of Keratin, type I cytoskeletal 13	-3.704	0.000
P13667	PDIA4	Protein disulfide-isomerase A4	-0.798	0.000
P13796	LCP1	Plastin-2	0.362	0.049
P07237	P4HB	Protein disulfide-isomerase	0.302	0.027
P48735	IDH2	Isocitrate dehydrogenase [NADP], mitochondrial	0.551	0.012
P23368	ME2	NAD-dependent malic enzyme, mitochondrial	Increased in CF	0.000

P07339	CTSD	Cathepsin D	-2.120	0.000
P07384	CAPN1	Calpain-1 catalytic subunit	-0.864	0.001
P23528	CFL1	Cofilin-1	-	n.s.
P41218	MNDA	Myeloid cell nuclear differentiation antigen	1.051	0.000
P31943	HNRNPH1	Heterogeneous nuclear ribonucleoprotein H	-	n.s.
P15924	DSP	Isoform DPI of Desmoplakin	-2.154	0.000
Q15080	NCF4	Neutrophil cytosol factor 4	Increased in CF	0.000
Q15149	PLEC	Plectin	-1.452	0.000
P09622	DLD	Dihydrolipoyl dehydrogenase, mitochondrial	-0.930	0.000
Q6P4A8	PLBD1	Phospholipase B-like 1	-	n.s.
P51149	RAB7A	Ras-related protein Rab-7a	-0.379	0.022
P51159	RAB27A	Ras-related protein Rab-27A	1.698	0.000
P00367	GLUD1	Glutamate dehydrogenase 1, mitochondrial	-1.568	0.000
P00403	MT-CO2	Cytochrome c oxidase subunit 2	-0.759	0.006
P00450	CP	Ceruloplasmin	-0.662	0.016
P43353	ALDH3B1	Aldehyde dehydrogenase family 3 member B1	-0.878	0.001
P00505	GOT2	Aspartate aminotransferase, mitochondrial	-1.137	0.000
P43490	NAMPT	Nicotinamide phosphoribosyltransferase	0.797	0.000
P61626	LYZ	Lysozyme C	-0.548	0.002
P35579	MYH9	Myosin-9	-	n.s.
P27797	CALR	Calreticulin	-	n.s.
P27824	CANX	Calnexin	-	n.s.
P10606	COX5B	Cytochrome c oxidase subunit 5B, mitochondrial	-	n.s.
P10644	PRKAR1A	cAMP-dependent protein kinase type I-alpha regulatory subunit	0.914	0.000
Q92542	NCSTN	Nicastrin	-	n.s.

P20160	AZU1	Azurocidin	1.693	0.000
Q6UX06	OLFM4	Olfactomedin-4	1.636	0.000
P55072	VCP	Transitional endoplasmic reticulum ATPase	-1.079	0.000
P55084	HADHB	Trifunctional enzyme subunit beta, mitochondrial	-1.211	0.000
P12273	PIP	Prolactin-inducible protein	-1.746	0.000
Q08380	LGALS3BP	Galectin-3-binding protein	-2.038	0.000
O75131	CPNE3	Copine-3	1.068	0.000
P30040	ERP29	Endoplasmic reticulum resident protein 29	1.499	0.000
P12724	RNASE3	Eosinophil cationic protein	Increased in CF	0.001
P60842	EIF4A1	Eukaryotic initiation factor 4A-I	Increased in CF	0.004
P04792	HSPB1	Heat shock protein beta-1	-3.468	0.000
P04843	RPN1	Dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit 1 precursor	-1.065	0.000
P14314	PRKCSH	Glucosidase 2 subunit beta	-	n.s.
P14625	HSP90B1	Endoplasmin	-0.391	0.009
P40121	CAPG	Macrophage-capping protein	-	n.s.
O75955	FLOT1	Flotillin-1	0.782	0.001
P30740	SERPINB1	Leukocyte elastase inhibitor	1.050	0.000
P14780	MMP9	Matrix metalloproteinase-9	-	n.s.
P22894	MMP8	Neutrophil collagenase	1.777	0.000
P08311	CTSG	Cathepsin G	2.147	0.000
P61158	ACTR3	Actin-related protein 3	-0.648	0.000
O60234	GMFG	Glia maturation factor gamma	1.663	0.000
P50395	GDI2	Rab GDP dissociation inhibitor beta	-	n.s.
P40939	HADHA	Trifunctional enzyme subunit alpha, mitochondrial	-1.428	0.000

P00558	PGK1	Phosphoglycerate kinase 1	-	n.s.
P13639	EEF2	Elongation factor 2	-1.224	0.000
P10599	TXN	Thioredoxin	-1.289	0.000
P07737	PFN1	Profilin-1	-	n.s.
P30038	ALDH4A1	Delta-1-pyrroline-5-carboxylate dehydrogenase, mitochondrial	-	n.s.
P04839	CYBB	Cytochrome b-245 heavy chain	1.197	0.000
P09917	ALOX5	Arachidonate 5-lipoxygenase	-	n.s.
P04083	ANXA1	Annexin A1	-1.957	0.000
P52209	PGD	6-phosphogluconate dehydrogenase, decarboxylating	-	n.s.
P27105	STOM	Erythrocyte band 7 integral membrane protein	1.685	0.000
P09211	GSTP1	Glutathione S-transferase P	-0.741	0.000
P15144	ANPEP	Aminopeptidase N	0.486	0.007
Q8WUM4	PDCD6IP	Programmed cell death 6-interacting protein	-	n.s.
Q12913	PTPRJ	Receptor-type tyrosine-protein phosphatase eta precursor	0.621	0.006
P40926	MDH2	Malate dehydrogenase, mitochondrial	-1.000	0.000
Q8TDL5	BPIFB1	BPI fold-containing family B member 1	-1.470	0.003
Q14739	LBR	Lamin-B receptor	-	n.s.
P49913	CAMP	Cathelicidin antimicrobial peptide precursor	1.215	0.002
Q8IX19	MCEMP1	Mast cell-expressed membrane protein 1	Increased in CF	0.000
Q9Y2J8	PADI2	Protein-arginine deiminase type-2	0.915	0.005
P39656	DDOST	Dolichyl-diphosphooligosaccharide--protein glycosyltransferase 48 kDa subunit	-0.909	0.001
Q13576	IQGAP2	Ras GTPase-activating-like protein IQGAP2	0.546	0.001
P80188	LCN2	Neutrophil gelatinase-associated lipocalin	1.806	0.000

Q15084	PDIA6	Protein disulfide-isomerase A6	-0.570	0.035
P21333	FLNA	Isoform 2 of Filamin-A	-0.897	0.000
P50990	CCT8	T-complex protein 1 subunit theta	-0.833	0.000
P06576	ATP5B	ATP synthase subunit beta, mitochondrial	-1.416	0.000
A6NI72	NCF1B	Putative neutrophil cytosol factor 1B	Increased in CF	0.000
P10909	CLU	Clusterin	-2.856	0.000
P69905	HBA1 & HBA2	Hemoglobin subunit alpha	-	n.s.
P07355	ANXA2	Isoform 2 of Annexin A2	-2.744	0.000
P62937	PPIA	Peptidyl-prolyl cis-trans isomerase A	-	n.s.
P25705	ATP5A1	ATP synthase subunit alpha, mitochondrial	-0.751	0.010
P62805	HIST1H4F/ HIST1H4C; HIST1H4A; HIST1H4I; HIST1H4B; HIST2H4B; HIST1H4L; HIST1H4E; HIST1H4D; HIST1H4J; HIST1H4K; HIST1H4H; HIST2H4A; HIST4H4	Histone H4	1.775	0.000
Q14204	DYNC1H1	Cytoplasmic dynein 1 heavy chain 1	Reduced in CF	0.001
P06733	ENO1	Alpha-enolase	-0.694	0.002
P04040	CAT	Catalase	1.108	0.000

P14618	PKM	Pyruvate kinase PKM	-	n.s.
P13489	RNH1	Ribonuclease inhibitor	-0.549	0.001
P01011	SERPINA3	Alpha-1-antichymotrypsin	0.548	0.048
P01009	SERPINA1	Alpha-1-antitrypsin	1.036	0.001
P29401	TKT	Transketolase	0.955	0.000
P22314	UBA1	Ubiquitin-like modifier-activating enzyme 1	-	n.s.
P37837	TALDO1	Transaldolase	-0.581	0.019
O75083	WDR1	WD repeat-containing protein 1	-	n.s.
P01024	C3	Complement C3	Reduced in CF	0.020
Q14254	FLOT2	Flotillin-2	-	n.s.
P50995	ANXA11	Annexin A11	1.198	0.000
Q5JTV8	TOR1AIP1	Torsin-1A-interacting protein 1	-0.511	0.001
O43451	MGAM	Maltase-glucoamylase, intestinal	2.495	0.000
P36222	CHI3L1	Chitinase-3-like protein 1	CF only	n.a.
P61160	ACTR2	Actin-related protein 2	CF only	n.a.
Q9H9B4	SFXN1	Sideroflexin-1	CF only	n.a.
P48595	SERPINB10	Serpin B10	CF only	n.a.
Q13043	STK4	Serine/threonine-protein kinase 4	CF only	n.a.
P15498	VAV1	Proto-oncogene vav	CF only	n.a.
Q13636	RAB31	Ras-related protein Rab-31	CF only	n.a.
Q9H4M9	EHD1	EH domain-containing protein 1	CF only	n.a.
Q99829	CPNE1	Copine-1	CF only	n.a.
P20592	MX2	Interferon-induced GTP-binding protein Mx2	CF only	n.a.
P57737	CORO7	Coronin-7	CF only	n.a.
P34059	GALNS	N-acetylgalactosamine-6-sulfatase	CF only	n.a.
P26038	MSN	Moesin	CF only	n.a.

P08473	MME	Neprilysin	CF only	n.a.
Q9P107	GMIP	GEM-interacting protein	CF only	n.a.
P07332	FES	Tyrosine-protein kinase Fes/Fps	CF only	n.a.
Q9NUQ9	FAM49B	Protein FAM49B	CF only	n.a.
P98171	ARHGAP4	Rho GTPase-activating protein 4	CF only	n.a.
Q86YV0	RASAL3	RAS protein activator like-3	CF only	n.a.
Q14005	IL16	Pro-interleukin-16	CF only	n.a.
P30153	PPP2R1A	Serine/threonine-protein phosphatase 2A 65 kDa regulatory subunit A alpha isoform	CF only	n.a.
P62834	RAP1A	Ras-related protein Rap-1A	control only	n.a.
Q07954	LRP1	Prolow-density lipoprotein receptor-related protein 1	control only	n.a.
P02786	TFRC	Transferrin receptor protein 1	control only	n.a.
Q9NZM1	MYOF	Myoferlin	control only	n.a.
P10253	GAA	Lysosomal alpha-glucosidase	control only	n.a.
P07858	CTSB	Cathepsin B	control only	n.a.
Q5JXB2	UBE2NL	Putative ubiquitin-conjugating enzyme E2 N-like	control only	n.a.
Q96TC7	RMDN3	Regulator of microtubule dynamics protein 3	control only	n.a.
P13686	ACP5	Tartrate-resistant acid phosphatase type 5	control only	n.a.
Q5TD94	RSPH4A	Radial spoke head protein 4 homolog A	control only	n.a.
Q5RHP9	ERICH3	Glutamate-rich protein 3	control only	n.a.
P18124	RPL7	60S ribosomal protein L7	control only	n.a.
Q9NQ38	SPINK5	Serine protease inhibitor Kazal-type 5	control only	n.a.
Q8N392	ARHGAP1	Rho GTPase-activating protein 18	control only	n.a.

	8			
Q13838	DDX39B	Spliceosome RNA helicase DDX39B	control only	n.a.
P19013	KRT4	Keratin, type II cytoskeletal 4	control only	n.a.
O15031	PLXNB2	Plexin-B2	control only	n.a.

Table E3: Proteins detected in all 48 sputum samples or all samples belonging to at least one cohort. Where independent samples T-tests identified differential abundance between the cohorts ($p<0.05$) the $\log_2$ ratio of mean abundance levels and significance are recorded. Where Mann-Whitney U Tests identified differential abundance ($p<0.05$) increased or decreased abundance in CF is recorded along with significance. No significant difference in abundance between the cohorts is recorded as "-".

Functions Category	Bio Functions Annotation	Confidence level (*p*-value)	Regulation z-score
Cell Death	apoptosis of neutrophils	3.56E-09	-1.308
Cell Death	apoptosis of myeloid cells	6.85E-07	-1.108
Cell Death	killing of cells	6.31E-05	1.88
Cell Death	cell survival	3.15E-04	-1.014
Cell Death	cell death	3.42E-08	0.875
Cell Death	apoptosis of phagocytes	9.75E-07	-0.913
Cell Death	apoptosis	2.90E-06	0.648
Cell Death	cell death of lung cell lines	2.03E-05	
Cell Death	cell viability	1.41E-04	-0.895
Cell Death	cell death of muscle cells	2.36E-04	0.56
Cell Death	cell death of organ	2.63E-04	0.885
Cell Death	apoptosis of organ	2.86E-04	0.29
Cell Death	apoptosis of lung cell lines	3.34E-04	
Cell Death	apoptosis of muscle cells	3.45E-04	0.342
Cell-To-Cell Signaling & Interaction	adhesion of granulocytes	5.07E-04	1.465
Cell-To-Cell Signaling & Interaction	activation of cells	2.63E-03	1.06
Cell-To-Cell Signaling & Interaction	response of granulocytes	4.84E-04	-0.392
Cellular Assembly & Organization	disruption of lipid bilayer	1.88E-04	
Cellular Assembly & Organization	fusion of liposome	2.82E-04	

Cellular Compromise	injury of cells	1.06E-04	-0.261
Cellular Function & Maintenance	cellular homeostasis	5.25E-04	1.165
Cellular Function & Maintenance	autophagy	3.62E-04	0.607
Cellular Growth & Proliferation	proliferation of immune cells	3.68E-03	-2.182
Cellular Growth & Proliferation	proliferation of muscle cells	6.08E-03	1.581
Cellular Growth & Proliferation	proliferation of lymphocytes	8.17E-03	-1.975
Cellular Growth & Proliferation	growth of cells	4.81E-06	-0.016
Cellular Movement	migration of cells	2.71E-07	1.311
Cellular Movement	leukocyte migration	3.86E-06	1.82
Cellular Movement	cell movement of myeloid cells	1.10E-05	1.68
Cellular Movement	cell movement of granulocytes	2.36E-05	1.121
Cellular Movement	chemotaxis of cells	5.03E-05	1.142
Cellular Movement	invasion of cells	1.44E-04	1.001
Cellular Movement	cell movement of antigen presenting cells	2.31E-04	1.793
Cellular Movement	recruitment of leukocytes	3.66E-03	1.188
Cellular Movement	cell movement	5.53E-07	0.987
Cellular Movement	migration of granulocytes	3.17E-04	0.578
Free Radical Scavenging	metabolism of reactive oxygen species	1.85E-05	1.655
Free Radical Scavenging	generation of reactive oxygen species	4.71E-05	1.534
Free Radical Scavenging	synthesis of reactive oxygen	6.23E-05	1.729

	species		
Inflammatory Response	chemotaxis of leukocytes	1.03E-05	1.426
Inflammatory Response	cell movement of neutrophils	2.25E-05	2.036
Inflammatory Response	inflammatory response	2.92E-05	1.652
Inflammatory Response	cell movement of phagocytes	7.17E-05	2.272
Inflammatory Response	chemotaxis of granulocytes	8.05E-05	2.135
Inflammatory Response	chemotaxis of myeloid cells	8.29E-05	2.339
Inflammatory Response	chemotaxis of phagocytes	1.04E-04	1.641
Inflammatory Response	cell movement of macrophages	1.97E-04	1.877
Inflammatory Response	migration of neutrophils	9.41E-04	1.476
Inflammatory Response	chemotaxis of neutrophils	2.42E-03	2.284
Inflammatory Response	activation of myeloid cells	4.00E-03	1.195
Inflammatory Response	cell movement of monocytes	4.92E-03	1.481
Inflammatory Response	immune response	2.81E-06	0.689
Inflammatory Response	phagocytosis of granulocytes	4.41E-04	
Organismal Injury & Abnormalities	injury of organ	1.72E-05	-0.021
Post-Translational Modification	N-glycosylation of protein	7.48E-05	
Protein Synthesis	degradation of protein	1.02E-06	-0.653
Protein Synthesis	metabolism of protein	4.39E-06	-0.732
Protein Synthesis	damage of protein	1.88E-04	
Small Molecule Biochemistry	production of nitric oxide	7.49E-03	-1.821
Small Molecule Biochemistry	catabolism of acidic amino acid	3.93E-04	

Table E4: Ingenuity Pathway Analysis Bio Functions predicted by Fisher's Exact Test as over-represented in CF (p<0.05) with Regulation z-scores >1.0 or <-1.0 (i.e. indicating a greater than 2-fold difference). Shaded rows record Bio Functions with a high prediction confidence (p<0.0005), but where the predicted Regulation z-score indicates a less than 2-fold difference or none could be calculated.

UniProt Accession	Gene Name	Description	GO Biological Processes
P36222	CHI3L1	Chitinase-3-like protein 1	activation of NF-kappaB-inducing kinase activity; carbohydrate metabolic process; inflammatory response
P61160	ACTR2	Actin-related protein 2	cellular component movement
Q9H9B4	SFXN1	Sideroflexin-1	erythrocyte differentiation; cellular iron ion homeostasis
P48595	SERPINB10	Serpin B10	regulation of proteolysis
Q13043	STK4	Serine/threonine-protein kinase 4	induction of apoptosis; negative regulation of cell proliferation
P15498	VAV1	Proto-oncogene vav	induction of apoptosis by extracellular signals; immune response; regulation of small GTPase mediated signal transduction
Q13636	RAB31	Ras-related protein Rab-31	Golgi vesicle transport; protein transport; small GTPase mediated signal transduction
Q9H4M9	EHD1	EH domain-containing protein 1	GTP catabolic process; endocytic recycling; intracellular protein transport
Q99829	CPNE1	Copine-1	lipid metabolic process; vesicle-mediated transport
P20592	MX2	Interferon-induced GTP-binding protein Mx2	regulation of cell cycle; regulation of nucleocytoplasmic transport
P57737	CORO7	Coronin-7	Golgi membrane; cytoplasmic membrane-bounded vesicle
P34059	GALNS	N-acetylgalactosamine-6-sulfatase	carbohydrate metabolic process; keratan sulfate catabolic process
P26038	MSN	Moesin	leukocyte cell-cell adhesion; leukocyte migration; regulation of lymphocyte migration
P08473	MME	Neprilysin	beta-amyloid metabolic process; cellular response to cytokine stimulus; proteolysis
Q9P107	GMIP	GEM-interacting protein	negative regulation of Rho GTPase activity
P07332	FES	Tyrosine-protein kinase Fes/Fps	regulation of cell proliferation; positive regulation of myeloid cell differentiation; regulation of mast cell degranulation
Q9NUQ9	FAM49B	Protein FAM49B	-
P98171	ARHGAP4	Rho GTPase-activating protein 4	Rho protein signal transduction; induction of apoptosis by extracellular signals; cytoskeleton organization
Q86YV0	RASAL3	RAS protein activator like-3	negative regulation of Ras protein signal transduction
Q14005	IL16	Pro-interleukin-16	immune response; induction of positive chemotaxis; leukocyte chemotaxis
P30153	PPP2R1A	Serine/threonine-protein phosphatase 2A 65 kDa regulatory subunit A alpha isoform	inactivation of MAPK activity; induction of apoptosis; negative regulation of cell growth

P23368	ME2	NAD-dependent malic enzyme, mitochondrial	malate metabolic process
Q15080	NCF4	Neutrophil cytosol factor 4	immune response; phagosome maturation; antigen processing and presentation of exogenous peptide antigen via MHC class I, TAP-dependent; cell communication
P12724	RNASE3	Eosinophil cationic protein	RNA catabolic process; defense response to bacterium
Q8IX19	MCEMP1	Mast cell-expressed membrane protein 1	-
A6NI72	NCF1B	Putative neutrophil cytosol factor 1B	cell communication
P62834	RAP1A	Ras-related protein Rap-1A	activation of MAPKK activity; energy reserve metabolic process; small GTPase mediated signal transduction
Q13838	DDX39B	Spliceosome RNA helicase DDX39B	ATP-dependent helicase activity
Q5JXB2	UBE2NL	Putative ubiquitin-conjugating enzyme E2 N-like	-
Q5TD94	RSPH4A	Radial spoke head protein 4 homolog A	cilium movement
Q9NQ38	SPINK5	Serine protease inhibitor Kazal-type 5	epithelial cell differentiation; negative regulation of immune response; negative regulation of proteolysis
P10253	GAA	Lysosomal alpha-glucosidase	glucose metabolic process; glycogen catabolic process; lysosome organization
Q96TC7	RMDN3	Regulator of microtubule dynamics protein 3	apoptotic process; cell differentiation
P19013	KRT4	Keratin, type II cytoskeletal 4	cytoskeleton organization; epithelial cell differentiation; negative regulation of epithelial cell proliferation
P07858	CTSB	Cathepsin B	autophagy; negative regulation of cell death; regulation of apoptotic process; proteolysis
Q5RHP9	ERICH3	Glutamate-rich protein 3	-
Q8N392	ARHGAP18	Rho GTPase-activating protein 18	positive regulation of GTPase activity
Q9NZM1	MYOF	Myoferlin	plasma membrane repair; regulation of vascular endothelial growth factor receptor signaling pathway
P02786	TFRC	Transferrin receptor protein 1	cellular iron ion homeostasis; proteolysis
P18124	RPL7	60S ribosomal protein L7	structural constituent of ribosome; transcription regulator activity; translation
O15031	PLXNB2	Plexin-B2	positive regulation of GTPase activity
P13686	ACP5	Tartrate-resistant acid phosphatase type 5	response to LPS; negative regulation of inflammatory response; negative regulation of nitric oxide biosynthetic process;

			negative regulation of superoxide anion generation
Q07954	LRP1	Prolow-density lipoprotein receptor-related protein 1	apoptotic cell clearance; beta-amyloid clearance; regulation of phospholipase A2 activity; regulation of actin cytoskeleton organization
Q14764	MVP	Major vault protein	cell proliferation; negative regulation of protein autophosphorylation; negative regulation of signaling; protein activation cascade; protein transport
Q14204	DYNC1H1	Cytoplasmic dynein 1 heavy chain 1	antigen processing and presentation of exogenous peptide antigen via MHC class II; cell death; microtubule-based movement; stress granule assembly
P01024	C3	Complement C3	complement activation, alternative & classical pathways; inflammatory response; negative regulation of endopeptidase activity; positive regulation of activation of membrane attack complex; positive regulation of angiogenesis; positive regulation of glucose transport; positive regulation of lipid storage; positive regulation of phagocytosis; regulation of triglyceride biosynthetic process

Table E5: Gene Ontology Biological Processes for proteins detected exclusively in the CF cohort (dark red) or the control cohort (dark blue) and proteins up-regulated (pale red) or down-regulated (pale blue) in CF relative to controls (Mann-Whitney U test; $p<0.05$).

Run	Samples
1	CF1.1, CF1.2, CF1.3, CF2.1, CF2.2, CF2.3 & REF
2	CF3.1, CF3.2, CF3.3, CF6.1, CF6.2, CF6.3 & REF
3	CF4.1, CF4.2, CF4.3, CF5.1, CF5.2, CF5.3 & REF
4	CF7.1, CF7.2, CF7.3, CF11.1, CF11.2, CF11.3 & REF
5	CF12.1, CF12.2, CF12.3, CF13.1, CF13.2, CF13.3 & REF
6	CF15.1, CF15.2, CF15.3, CF16.1, CF16.2, CF16.3 & REF
7	HC23, HC24, HC33, HC35, HC36, HC38 & REF
8	HC27, HC28, HC29, HC39, HC40, HC41 & REF

Table E6: Samples included per MudPIT run. CF indicates the CF cohort and HC the healthy control cohort. The code used to label the CF samples indicates first patient ID followed by ".1" for exacerbation pre-antibiotic treatment, ".2" for exacerbation post-antibiotic treatment (<24h) and ".3" for stable follow-up (4-15 weeks following treatment.) Only one sample was collected from each of the health controls.

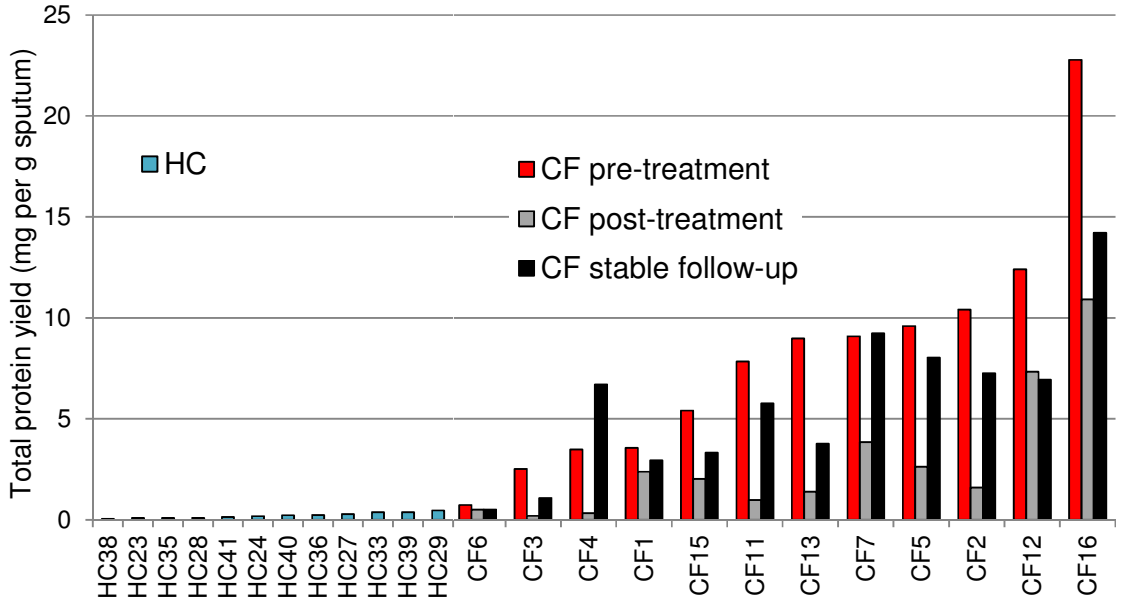


Figure E1: Sputum total protein yields (mg per g sputum) for single samples collected from healthy controls (HC) and for three samples collected from each individual in the CF cohort. The three CF samples were collected: (1) at exacerbation before commencement of antibiotic treatment; (2) at completion of antibiotic treatment for exacerbation; and (3) during a follow-up period of clinical stability.

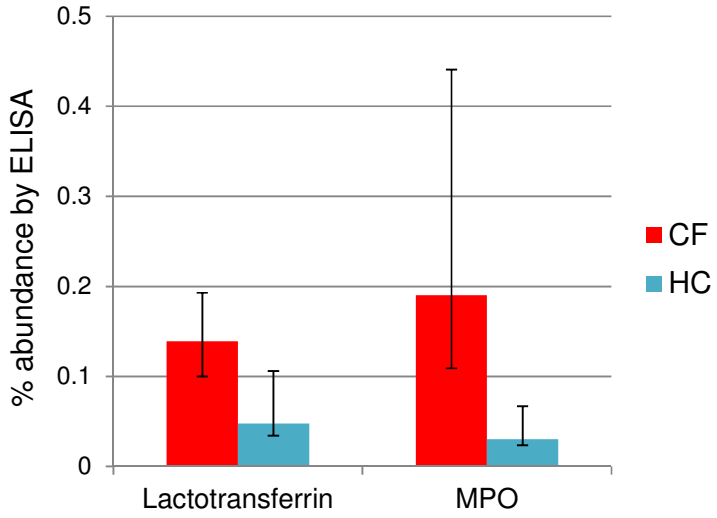


Figure E2: Median (±IQR) % abundance of myeloperoxidase and lactotransferrin per sample as determined by ELISA quantification. For both proteins, median % abundance is greater in the CF cohort compared to the healthy control (HC) cohort ($p<0.05$, Mann Whitney U test).

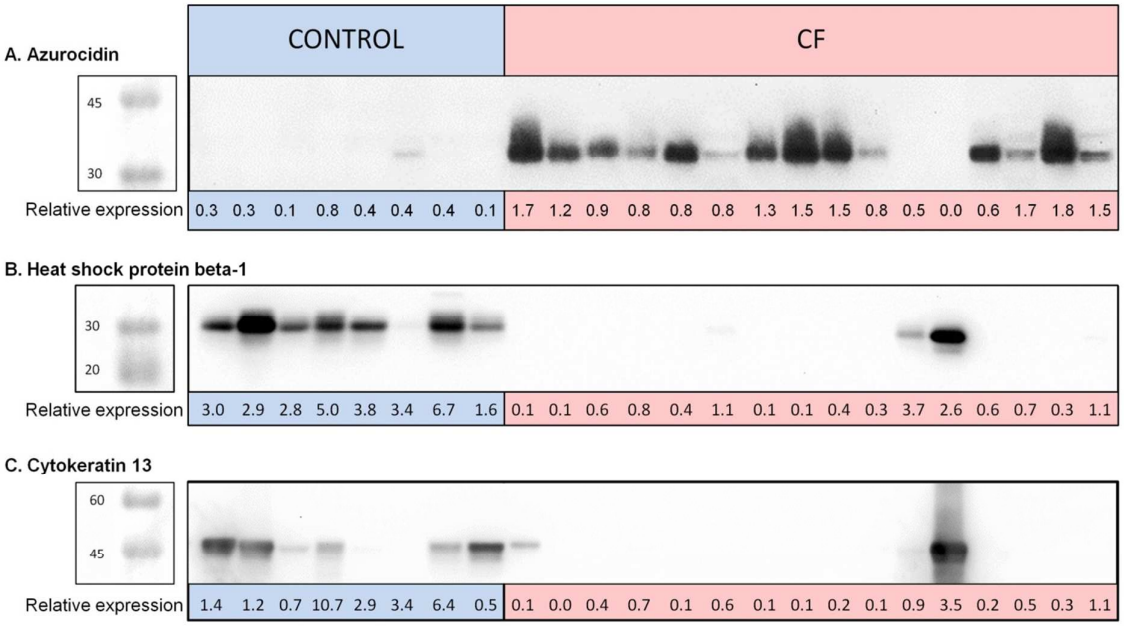


Figure E3: Western Blot validation of mass spectrometry data for azurocidin, heat shock protein beta-1 and cytokeratin 13. Relative abundance values were calculated as % emPAI from the mass spectrometry data.

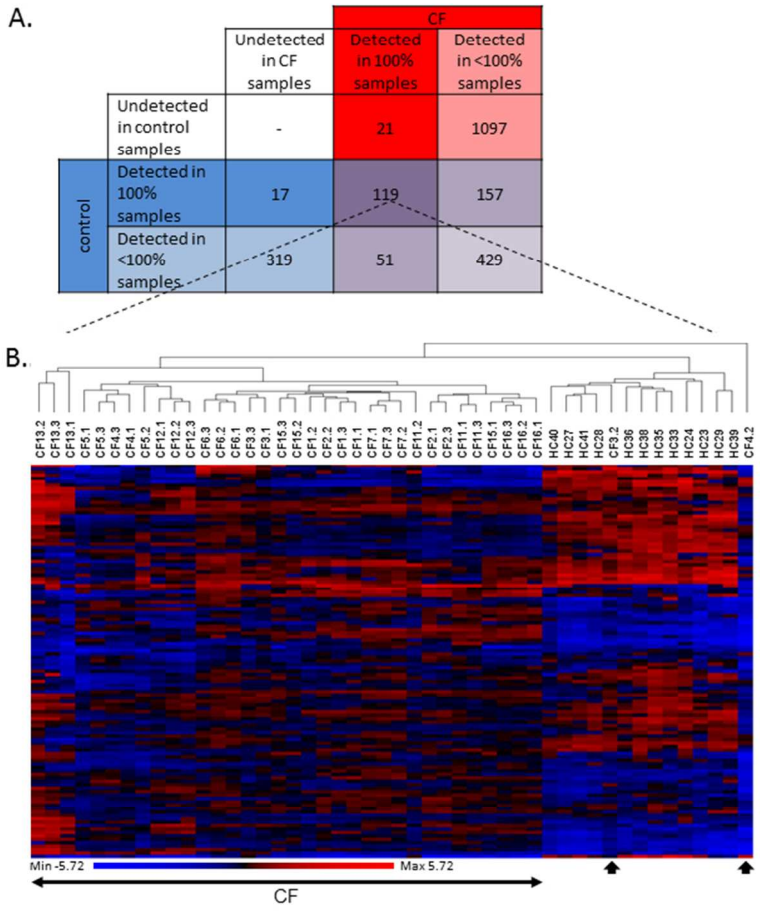


Figure 1: Distribution and relative abundance of human proteins detected in sputum samples collected from CF and healthy control cohorts. A. shows that 119 proteins were found in all 48 samples, i.e. both CF and control, while a further 21 proteins were found exclusively in all CF samples and 17 proteins exclusively in all control samples. B. shows hierarchical clustering (using Euclidean distance dissimilarity and McQuitty's linkage criteria) of individual sputum samples according to the relative abundance levels for all 119 consistently detected proteins. CF and control samples largely cluster separately with the only exceptions being 2 CF samples (marked with black arrows) collected after antibiotic treatment for pulmonary exacerbation. The scale is a log2 ratio of relative abundance.

190x254mm (96 x 96 DPI)

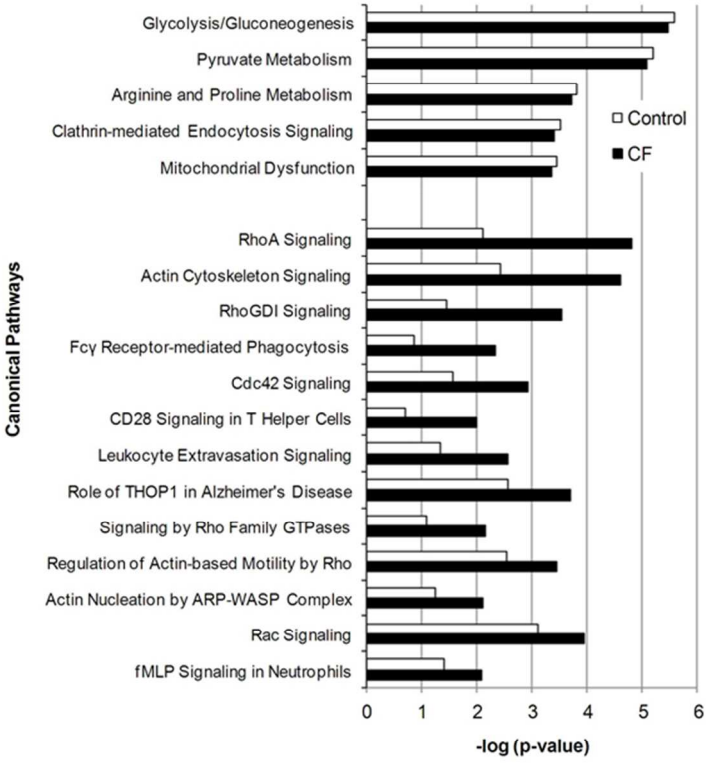


Figure 2: Differential activity of Canonical Pathways in consensus CF and control proteomes. Based on the set of proteins detected in all samples within a cohort, Ingenuity Pathway Analysis software applied Fisher's Exact Test to predict the Canonical Pathways likely to be active within that cohort ($p<0.05$). The top 5 pathways predicted in the control cohort concern general metabolism, clatherin-mediated endocytosis signaling and mitochondrial dysfunction and these were detected with approximately equal probabilities in both cohorts. Pathways where there is a >5-fold increase in p-value in the CF proteome compared to the control proteome, indicating likely up-regulation in CF are involved in immune functions including neutrophil recruitment, rearrangement of the actin cytoskeleton, phagocytosis and T cell signaling.

190x254mm (96 x 96 DPI)

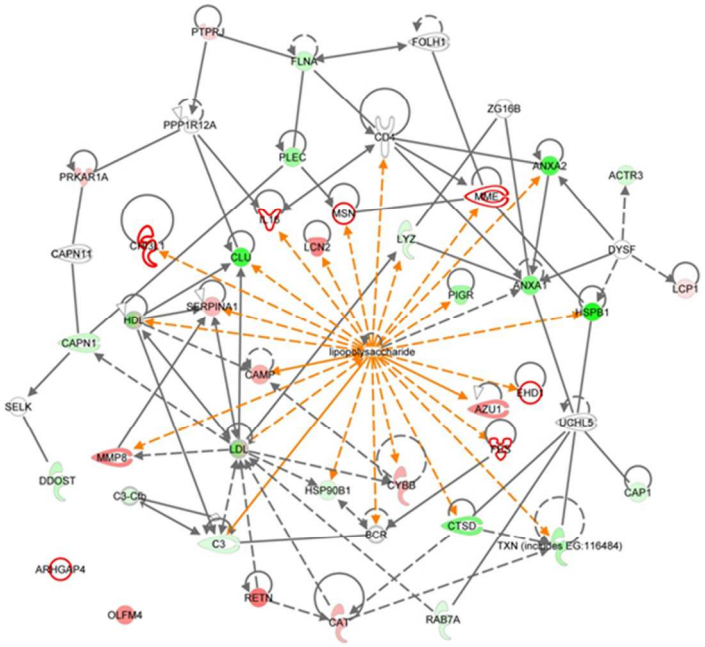
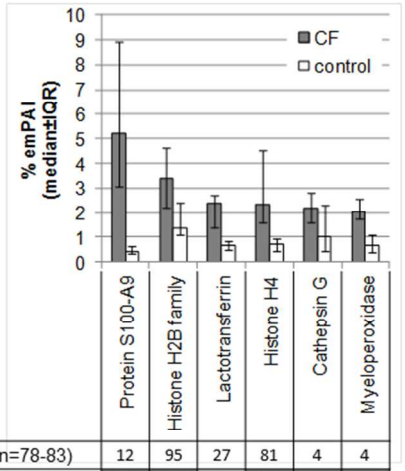


Figure 3: Ingenuity Pathway Analysis predicted network for proteins detected in CF involved in Inflammatory Response and Cell Movement Bio Functions. The intensity of red or green shading indicates the degree of increased or decreased abundance in CF relative to healthy controls, respectively. A red outline with no internal shading indicates that the protein was detected exclusively in the CF cohort. The network identifies LPS as a likely inflammatory trigger.

190x254mm (96 x 96 DPI)



Tissue Location	Protein S100-A9	Histone H2B family	Lactotransferrin	Histone H4	Cathepsin G	Myeloperoxidase
% tissues where expressed (n=78-83)	12	95	27	81	4	4
Bone marrow hematopoietic cells	H	H	M	H	H	H
Lung macrophages	M	M	L	L		M
Bronchus respiratory epithelial cells		H	L	H		
Nasopharynx respiratory epithelial cells		H	M	H		
Lung pneumocytes	L	M	L	M		

Figure 4: Relative abundance estimates (% emPAI median ± IQR) for proteins comprising >1% of total human proteins in >80% of CF samples. Human Protein Atlas (26) annotations for protein expression in terms of tissue location and localised abundance are recorded. H, M and L indicate high, medium and low protein expression. All proteins are either highly or moderately expressed in bone marrow hematopoietic cells, 50% of which are myelopoietic cells, and all except cathepsin G are moderately or lowly expressed in lung macrophages.

190x254mm (96 x 96 DPI)